Summer 2003

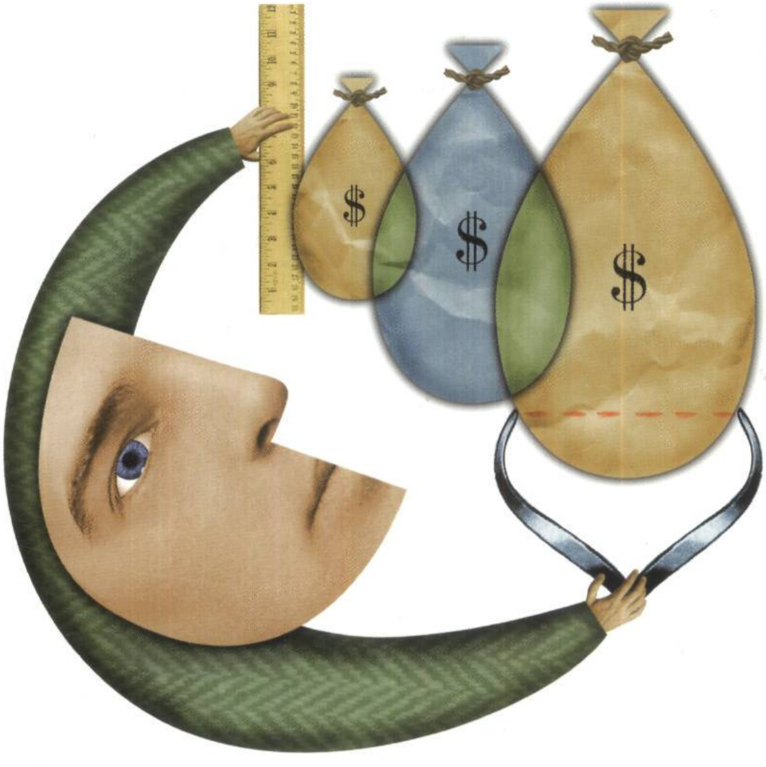
EDUCATION next

A JOURNAL OF OPINION AND RESEARCH

Are Teachers Underpaid?

New facts on an old debate

PETER TEMIN • RICHARD VEDDER • MICHAEL PODGURSKY



IN THIS ISSUE:

New Evidence on Accountability, An Inside Look at Discipline, Shanker's Strategy, Stuyvesant High School, School Choice Lockdown, Class-Size Reduction

CONTRIBUTORS: RONALD BROWNSTEIN, THOMAS DEE, MARC EPSTEIN, RICHARD KAHLENBERG, ERIC HANUSHEK, JENS LUDWIG, MARGARET RAYMOND, SOL STERN, MARTIN WEST, LUDGER WOESSMANN

www.educationnext.org / $7.00

AEI

1150 Seventeenth Street, N.W.
Washington, DC 20036
202.862.5800
www.aei.org

Education Policy Program
at the American Enterprise Institute

The American Enterprise Institute's work on education policy aims to improve the quality of education that children receive in America by rethinking the content, delivery, and governance of schooling. Through careful empirical research, open discussions, and thoughtful publications and conferences, AEI emphasizes sound reasoning over current popular trends and partisan politics. The members of AEI's Education Policy Program offer both policy criticisms and recommendations to enrich the national education debate. ***They tackle many of the most controversial questions*** of education policy, unflinchingly addressing issues that include school choice, racial inequality, teacher quality, curricular standards, accountability, higher education, civic values, and youth culture.

Douglas J. Besharov, *Joseph J. and Violet Jacobs Scholar in Social Welfare Studies*
Author of *Recognizing Child Abuse* (1990) and editor of *America's Disconnected Youth: Toward a Preventative Strategy* (1999). He is currently involved in a multiyear study of career and technical education.

Lynne V. Cheney, *senior fellow*
Author of *America: A Patriotic Primer* (2002) and *Telling the Truth* (1996). She is the former director of the National Endowment for the Humanities and currently focuses on issues of education and culture.

Frederick M. Hess, *resident scholar*
Author of *Revolution at the Margins: The Impact of Competition on Urban School Systems* (2002) and *Spinning Wheels: The Politics of Urban School Reform* (1999). He is the executive editor of *Education Next* and his writing on education has appeared in publications including *Phi Delta Kappan, Teachers College Record, Social Science Quarterly, Urban Affairs Review, Educational Policy,* and *Education Week.*

Christina Hoff Sommers, *resident scholar*
Author of *The War Against Boys: How Misguided Feminism is Harming Our Young Men* (2000). Her writing has appeared in a wide range of publications, including the *Wall Street Journal,* the *New England Journal of Medicine,* and the *Journal of Philosophy.*

UPCOMING EDUCATION CONFERENCES AT AEI

October 23–24, 2003: **A Qualified Teacher in Every Classroom? Appraising Old Answers and New Ideas**
Supported by the National Research Initiative and the Spencer Foundation.

January 2004: **Leaving No Child Behind? Options for Kids in Failing Schools**
Supported by the Annie E. Casey Foundation and the Walton Family Foundation.
Presented in partnership with the Thomas B. Fordham Institute.

For more information about AEI's education policy program, go to http://www.aei.org/education.

ELECTRONIC REPRODUCTION PROHIBITED

VOLUME 3, NUMBER 3
EDUCATION next
A JOURNAL OF OPINION AND RESEARCH

7

34

40

forum

7 Are Teachers Underpaid?
A question of value
Low Pay, Low Quality by PETER TEMIN *Comparable Worth* by RICHARD VEDDER

features

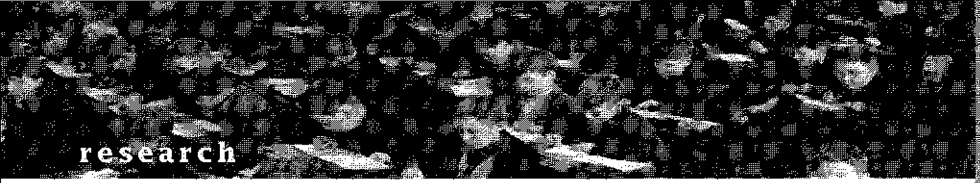
20 Façade of Excellence
Stuyvesant High School's dirty little secret by SOL STERN

28 Security Detail
An inside look at school discipline by MARC A. EPSTEIN

34 Philosopher or King?
The legacy of Albert Shanker by RICHARD D. KAHLENBERG

40 Locked Down
Will failing public schools let students leave? by RONALD BROWNSTEIN

48 High-Stakes Research
Accountability works after all by MARGARET E. RAYMOND & ERIC A. HANUSHEK



research

56 **Crowd Control**
Does reducing class size work?
by MARTIN R. WEST & LUDGER WOESSMANN

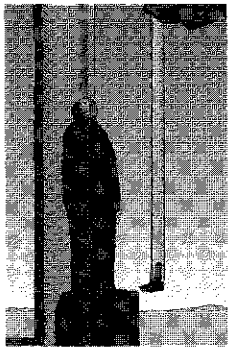
64 **Learning to Earn**
How higher standards affect graduation and employment
by THOMAS S. DEE

56

check the facts

71 **Fringe Benefits**
AFT and NEA teacher salary surveys
CHECKED by MICHAEL PODGURSKY

79 **Educational Achievement and Black-White Inequality**
The Great Unknown
by Jonathan Jacobsen et al.
CHECKED by JENS LUDWIG



79

from the editors

3 More Than Just Pay

correspondence

4 *Paper tiger unions; Bronze Age of education; and nimble charters*

book reviews

83 **Marginal Impact**
The small-schools literature REVIEWED by ROBIN J. LAKE

86 **Why Schools Matter: A Cross-National Comparison of Curriculum and Learning**
by William H. Schmidt et al. REVIEWED by STEPHEN P. HEYNEMAN

education matters to me

88 **Meet Mr. Shannon**
by KRISTIN KEARNS JORDAN

88

COVER ILLUSTRATION *by Joyce Hesselberth*

VOLUME 3, NUMBER 3

EDUCATION next

A JOURNAL OF OPINION AND RESEARCH

Editor-in-Chief	Paul E. Peterson
Senior Editor	Chester E. Finn Jr.
Executive Editors	Frederick M. Hess Marci Kanstoroom
Research Editor	Martin West
Articles Editor	Tyce Palmaffy
Managing Editor	Carol Peterson
Design Director	John Sizing
Graphics Editor	Tammy Frisby
Consulting Editors	Patricia Baker, Jay P. Greene
Editorial Staff	Susan Pasternack, Deborah Styles
Web Editor	Bridgitt Ahern
Circulation	Jennifer Presley
Editorial Board	John E. Chubb, Williamson Evers, Eric A. Hanushek, Paul Hill, E. D. Hirsch Jr., Caroline M. Hoxby, Terry M. Moe, Diane Ravitch, Herbert Walberg

Publisher	Hoover Institution, Leland Stanford Junior University

Additional Sponsoring Institutions	Program on Education Policy and Governance, Harvard University Thomas B. Fordham Foundation Manhattan Institute for Policy Research

Subscriptions One-year subscription (four issues), $20 in the United States; $30 in Canada; $40 elsewhere. To subscribe or to submit change of address information, call us at (800) 935-2882; fax us at (650) 723-8626; visit us at www.educationnext.org; or email us at hooverpress@hoover.stanford.edu. You may also write to *Education Next* Subscriptions, Hoover Institution, Stanford University, Stanford, CA 94305-6010. Cover price is $7/issue. ISSN 1539-9672.

Advertising For information on advertising, contact Advertising Manager, Program on Education Policy and Governance, Kennedy School of Government, Harvard University, Cambridge, MA 02138, (617) 495-6954.

Letters We welcome letters about our contents. Longer versions of some letters may be found at www.educationnext.org. Write to Letters to the Editor, *Education Next*, 226 Littauer North Yard, 1875 Cambridge Street, Cambridge, MA 02138; fax to (617) 496-1507; or email Editor_EM@latte.harvard.edu.

Editorial Queries We consider unsolicited contributions. Send manuscript submissions as email attachments to Editor_EM@latte.harvard.edu. Pieces sent by regular mail go to Carol Peterson, *Education Next*, 226 Littauer North Yard, 1875 Cambridge Street, Cambridge, MA 02138.

Opinions expressed in *Education Next* are those of the authors and do not necessarily reflect those of the Editorial Board of *Education Next*, the Board of Trustees of Leland Stanford Junior University, or the sponsors of *Education Next*.

Visit us at http://www.educationnext.org, where unabridged essays can be found.

© by the Board of Trustees of Leland Stanford Junior University, 2003

from the editors

More Than Just Pay

Higher salaries will accomplish little without bolder reforms

Are teachers paid too little? This has been more of an assumption than a question in recent decades. Even raising the issue carries the risk of being labeled anti-education. Nonetheless, a few of our intrepid authors dared to tackle the subject in this issue's pages. They suggest that today's teachers may be well paid, but that the nation may be undervaluing tomorrow's teachers.

Today's teachers are well paid in the sense that recent salary increases have outpaced inflation and stayed in step with the salaries of other educated professionals. So says **Richard Vedder** in this issue's forum. Vedder's claims are backed up by **Michael Podgursky**, who applies a critical eye to the surveys of teachers' salaries reported in U.S. Department of Education statistical publications.

> Teachers' salaries have stayed in step with those of other educated professionals.

However, the nation may be undervaluing tomorrow's teachers, argues **Peter Temin**, if we seek the high-caliber workforce of K–12 professionals that the late Albert Shanker envisioned during his decades as president of the American Federation of Teachers (see **Richard Kahlenberg**'s thoughtful retrospective on Shanker's contributions to American education). No longer can schools rely on the fact that women have few other occupational choices, or the fact that highly skilled workers are earning more in a technology-driven economy. **Martin West** and **Ludger Woessmann**'s analysis of data from 18 nations even suggests that higher pay for teachers might be more effective than reducing class size. The only significant benefits from class-size reduction were found in Greece and Iceland, where teachers are paid relatively low wages and are less educated.

However, simply paying teachers more will yield few benefits unless the recruitment and compensation of teachers are freed from constraints that reward mediocrity. Three snapshots of school life in New York City reveal the degree to which bureaucratic rule-making has undermined excellence in urban schools, especially among teachers. As a parent, **Sol Stern** was shocked to discover that New York City's teachers' contract sets conditions that, in effect, require school leaders to favor seniority over skills and knowledge in the hiring of teachers at Stuyvesant High, one of the city's premier high schools.

Byzantine licensing rules cause headaches at Bronx Preparatory, one of the city's most exciting charter school ventures. In our back-page essay, school founder **Kristin Kearns Jordan** bemoans the precious time and money her teachers must waste fulfilling the state's mindless course requirements.

Marc Epstein's tale from inside Jamaica High School is the most disturbing. Here bizarre interpretations of judicial decisions have weakened school authority, giving the worst elements within the school undue influence over the lives of teachers and students alike. No wonder urban schools have done so little to close the black-white test-score gap, a topic that **Jens Ludwig** deconstructs in this issue's check the facts.

Of course, students could just leave disorderly schools for better ones, a right guaranteed by the No Child Left Behind Act. But will the schools let them? Not likely, says **Ronald Brownstein.** Half-hearted efforts at compliance among urban school districts have kept the exodus to just a trickle.

Will the accountability provisions of No Child Left Behind prove more effective? **Margaret Raymond** and **Eric Hanushek** investigated state-level trends on the National Assessment of Educational Progress and found that the largest gains were among states that adopted accountability systems during the past decade. The authors' results differ sharply from those generated by a poorly designed study recently released to embarrassingly extensive media coverage.

Thomas Dee looks beyond test scores to examine the effects of accountability on high-school graduation rates and students' employment prospects. Using state-of-the-art techniques, Dee finds that earlier standards-based reforms lowered graduation rates but improved students' employability, especially among black students.

Taken together, these carry a broad message: If sensible disciplinary, choice, and accountability plans were combined with appropriate policies to recruit and compensate teachers, higher teacher pay could yield dividends for students, too.

–THE EDITORS

MISSION STATEMENT In the stormy seas of school reform, this journal will steer a steady course, presenting the facts as best they can be determined, giving voice (without fear or favor) to worthy research, sound ideas, and responsible arguments. Bold change is needed in American K–12 education, but *Education Next* partakes of no program, campaign, or ideology. It goes where the evidence points.

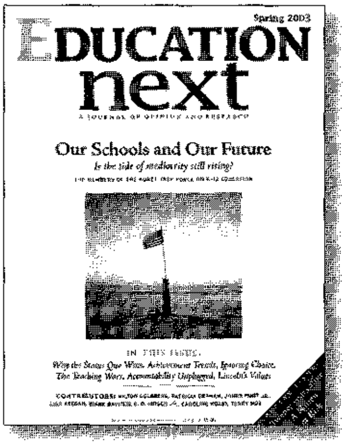
Paper tigers

Terry Moe ("Reform Blockers," *Feature*, Spring 2003) writes with his usual analytic elegance and passionate commitment to reform. The only problem is that many of his arguments conflict with the facts. I am no fan of teacher unions, for the same reason that Moe articulates—they too often support the status quo. But Moe's tendency to regard them as all-powerful does not accord with education politics in the past decade.

Exhibit A is the No Child Left Behind Act (NCLB). In this case, most congressional Democrats supported a bill that was largely crafted by a Republican presidential administration. So which is it? Were the Democrats engaged in a collective electoral suicide pact? Did they not fear the unions' ability to punish them in the next election? Or did their traditional union allies actually support the law's accountability provisions? None of these scenarios supports Moe's storyline.

An alternative explanation is that NCLB was all a big sham—the Democrats (and Republicans) pretended that they were seeking real reform; the unions pretended to oppose this pretense of reform; and governors and school administrators will pretend to implement this pretense of reform. This would accord with Moe's view of education politics. Yet it seems convoluted beyond plausibility. NCLB contains real provisions that will complicate the lives of public school teachers and administrators for years to come. If teacher unions are truly antireform, they *lost* on this law.

Of course, teacher unions act to protect their members; so do associations of lawyers, doctors, college professors, hairdressers, and undertakers. However, there is nothing inconsistent in wanting both to enhance one's job *and* to help the people with whom one works. Does Moe believe that police officers are indifferent to the murder rate in their city? Will firemen stand by while buildings burn down? Will tenured professors at community colleges or state universities shortchange their students as much as they can get away with? Surely there are some timeservers within all these groups, just as there are in teaching. Yet for the most part they are conscientious, hard-working professionals. We ought to assume that the vast majority of teachers also care deeply about their charges.


Spring 2003
EDUCATION next
Our Schools and Our Future
Is the tide of mediocrity still rising?
IN THIS ISSUE:
Why the Status Quo Wins, Achievement Trends, Ignoring Choice, The Teaching Wars, Accountability Unplugged, Lincoln's Values

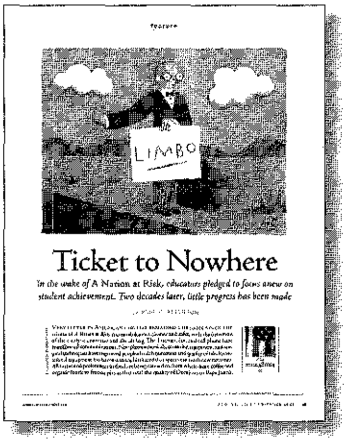
Moe is correct that core constituents of the Democratic party—"poor and minority families" in cities—endorse choice more enthusiastically than do Democratic elites. (It is worth noting that poor and minority families also endorse lots of other reforms, such as access to decent health care, to which Democratic and Republican elites fail to respond. School choice is not usually anywhere near the top of their list.) However, Moe does not point out that core constituents of the Republican party—middle-class and white families in suburbs—also disagree with their party elites on choice. Just about the last thing they want is choice programs that would open their schools to the children in failing urban schools. Both parties face a radical disconnect between the preferences of their elites and those of their voters, which may explain why fewer than 0.1 percent of the nation's school-age children are in publicly funded voucher programs.

JENNIFER HOCHSCHILD
Harvard University
Cambridge, Massachusetts

Bronze Age, at best

Let's begin with Paul Peterson's premise ("Ticket to Nowhere," *Feature*, Spring 2003) that the schools haven't matched America's economic and social progress in the past generation. The *A Nation at Risk* report warned in the starkest terms that if U.S. schools didn't shape up, the Germans and Japanese would beat our economic brains out, just as all the post-Sputnik Cassandras of the late 1950s and early 1960s warned that the Russians would win the cold war if the schools did not improve. Those children of the 1950s and 1960s are the ones who produced all the progress that Peterson now celebrates. Meanwhile, we've left the Germans, the Japanese, and the Russians in the dust.

The premise of progress is also partially wrong, in a way that illustrates the problems of the schools. For instance, America's improvements in health care have been horribly distributed, with a large percentage of Americans having no health coverage at all—just as the poorest Americans are generally forced to send their children to the worst schools. Most of the nations with which the United States is compared on international tests have far broader social-service networks, especially for children. It's wonderful to see conservatives using welfare states like Sweden, Denmark, Iceland, and France as benchmarks. In the United States, it's the teachers and



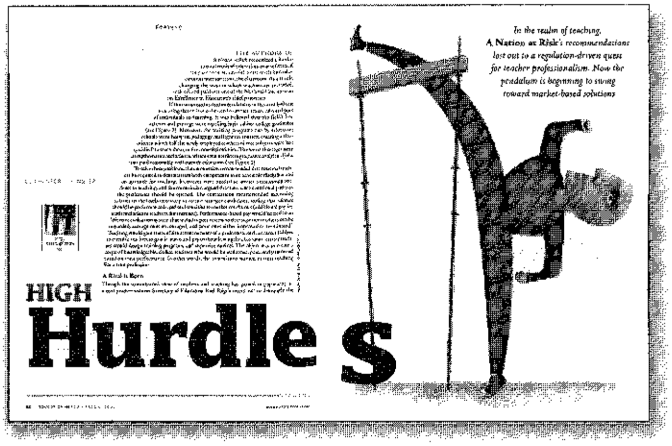
feature
Ticket to Nowhere
In the wake of A Nation at Risk, educators pledged to focus anew on student achievement. Two decades later, little progress has been made

schools that, for the most part, are the front line in dealing with the huge gaps in the social-service system.

Yes, America's schools are often outscored by other nations on international exams, but to what extent do those comparisons measure schools and to what extent are they reflections of cultural attitudes over which schools have little control? This country offers perpetual second chances and open access to some form of postsecondary education for virtually every student. Societies where high-stakes decisions are made by the age of 13, or even 18, concentrate students' and parents' minds wonderfully at a time when Americans, with their anti-intellectual traditions, are often thinking a lot more about Friday night football, cheerleader camps, drugs, and safe sex. It was not the schools that Ross Perot had to fight in the 1980s when he pushed his proposal that no student with less than a "C" average could play football; it was the local Texas football boosters.

Yes, American parents are better educated and wealthier than ever before. But the demographics don't tell us much about the conditions of parents' lives, and even less about those of their kids. What we do know is that after hitting a low of 14.9 percent in 1970, the percentage of children living in poverty hovered at somewhere between 19 and 22 percent through the 1980s and early 1990s. It dipped to 15.4 percent at the end of the Clinton years—still higher than 30 years before and considerably higher than in all those European countries. By now, it's almost certainly up again. Likewise, the percentage of kids who are not native English speakers is far higher now than it has been in nearly a century.

American schools, and particularly schools serving our poorest kids, should be far better, freer of stupid bureaucracies and self-serving teacher unions, staffed by better people, and supported by systems that encourage quality instead of fighting it. But 40 years of observing and writing about schools have also taught me how complex this story is, how interlocked schools are with their communities. We should be doing better, but it's not surprising that every generation, going back at least a century, has complained that the schools are failing and aren't as good as they used to be. Our schools were never as good as they used to be.



HIGH Hurdles

In the realm of teaching, A Nation at Risk's recommendations lost out to a regulation-driven quest for teacher professionalism. Now the pendulum is beginning to swing toward market-based solutions

PETER SCHRAG
The Sacramento Bee
Oakland, California

Demand side

In the area of teaching (Chester E. Finn Jr., "High Hurdles," *Feature*, Spring 2003), *A Nation at Risk* had it right. Potential teachers should demonstrate competence in an academic discipline and an aptitude for teaching; schools must offer incentives to attract outstanding candidates; unconventional paths to the profession must be forged; and salaries must be based on performance and sensitive to market conditions. Is there any hope that the next 20 years will see these reforms embedded in public education?

I am optimistic. Not because I think the establishment will willingly accept deregulation. Not because I believe multiple approaches will be tried under proper experimental conditions, providing research-based evidence for policymakers. I have no hope for a truce in the "teacher wars." Research will make a powerful contribution, but partisans will argue endlessly over methodology, and since when has a deep political divide been bridged by research? Chester Finn's plea for multiple approaches will, I fear, fall on deaf ears.

Nonetheless, I believe that performance will trump the "professionalism" movement. The market will demand it. Effective teachers in science, mathematics, special education, and other areas will continue to be in short supply.

Education is labor intensive. Urban districts typically spend up to 85 percent of their budgets on payroll. A labor-intensive business that cannot manage its workforce cannot perform at high levels. Voters are not going to back away from their demand for student achievement. Without flexible human-resource management policies, school districts are not going to be able to deliver high performance.

DONALD R. MCADAMS
Center for Reform of School Systems
Houston, Texas

Building charters

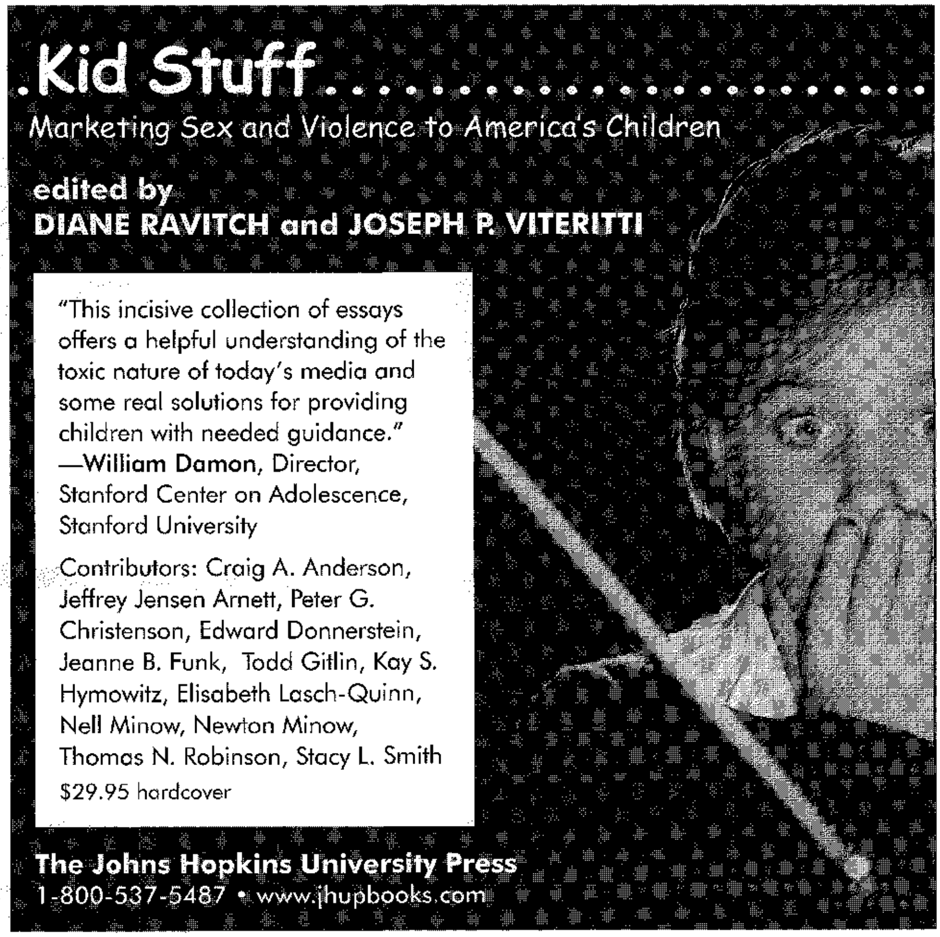
Bryan C. Hassel and Bruno M. Manno ("Charters Beset," *Forum*, Winter 2003) accurately describe the reform landscape as well as the challenges and opportunities for charter schools. Now entering their second decade, charter schools can clearly be considered a disruptive technology. There is plenty of evidence that school districts, service providers, policymakers, and parents are paying far more attention to charters than the U.S. Postal Service paid to Federal Express at a similar stage of development.

However, will these new schools significantly increase achievement, create systemic change, and stimulate innovative practices? Yes, but only if several things occur. First, charter schools must focus on student achievement and show their stuff on states' high-stakes exams. At a time of high-profile charter failures and evidence of fiscal mismanagement in a few charters, the movement can no longer fly under the flags of choice or competition. Charters must prove their academic value.

The current landscape, with rare exceptions, is filled with school districts that do not want to grant charters. Would Wal-Mart grant Target a charter to open a store across the street? Going forward, states must provide multiple authorizers. Let states, local districts, selected universities, and large-city mayors grant charters. Build the authorizers' capacity to champion their charters and ensure their success. Assist with a charter's ability to navigate the uncharted waters of facilities, permitting, and contracting. Consider what would happen if the U.S. Department of Education could grant charters in every state in the nation.

If charter schools are to increase achievement through competitive forces, that theory must be pilot tested. Few neighborhoods, if any, have multiple charter and noncharter public school choices. What would happen if a big-city mayor had the ability to bring six to ten independent charter schools to the same neighborhood? Would the proliferation of public school choices lead to increased achievement? Doing this in ten cities across the country would create a fantastic experiment.

Don Shalvey
CEO, *Aspire Public Schools*
Redwood City, California


Kid Stuff

Marketing Sex and Violence to America's Children

edited by
DIANE RAVITCH and JOSEPH P. VITERITTI

"This incisive collection of essays offers a helpful understanding of the toxic nature of today's media and some real solutions for providing children with needed guidance."
—**William Damon**, Director, Stanford Center on Adolescence, Stanford University

Contributors: Craig A. Anderson, Jeffrey Jensen Arnett, Peter G. Christenson, Edward Donnerstein, Jeanne B. Funk, Todd Gitlin, Kay S. Hymowitz, Elisabeth Lasch-Quinn, Nell Minow, Newton Minow, Thomas N. Robinson, Stacy L. Smith

$29.95 hardcover

The Johns Hopkins University Press
1-800-537-5487 • www.jhupbooks.com

The nimble charter

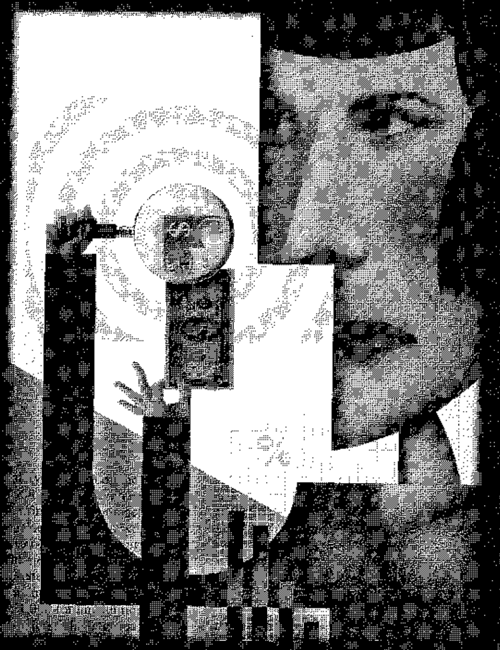
Bryan C. Hassel ("Friendly Competition," *Forum*, Winter 2003) is one of the finest scholars on charter schools, but he may overstate the organizational advantages of for-profit education management organizations (EMOs).

Hassel suggests that EMOs benefit from economies of scale. This is one of the biggest myths going around. First, two-thirds of every school budget is salaries, and no economies of scale are available there. In fact, in a nightmare scenario, Edison Schools may one day find itself organized by the union.

Second, I would guarantee that the overall per-pupil costs for goods and services at our tiny charter school are no greater than those of the EMOs. At the moment we're receiving $15,000 worth of pro-bono legal work for a real estate matter. Do you think law firms are donating to Edison? Sure, Edison can buy 1,000 Dell laptops and negotiate the price, but we can glom onto the Massachusetts state contract for public schools and enjoy even more purchasing power. Like the school district, the EMOs' ability to purchase in bulk is far less powerful than our ability to be nimble and quick.

Our "mom and pop" school probably shops more aggressively than individual Edison schools for cleaning contracts, and we probably require less cleaning since part of our school day involves kids' doing chores like emptying the trash. At the end of the year, I suspect we're better at getting every textbook returned in usable condition, since we're small enough to keep an eye on those things.

Michael Goldstein
The Match School
Boston, Massachusetts



Are Teachers Underpaid?

It is often said that the nation does not properly value teaching. Is this true?

AS THE NATION CONTINUES TO search for ways to upgrade the public school system, much attention has focused on how to improve the quality of K-12 teachers. The focus makes perfect sense: nothing in education is more important than what happens once the classroom door closes. The question is how to draw more high-caliber candidates to the profession. Raising salaries is one option, but it depends on whether one thinks teachers actually *are* underpaid. Our authors examine this widely held assumption.

ILLUSTRATIONS BY GORDON STUDER

Peter Temin finds a rare market failure **PAGE 8** Richard Vedder wants pay to respond to market forces **PAGE 14**

LOW PAY, low quality

For decades the nation has been able to school its children on the cheap by exploiting a trapped workforce of educated women.

Those days are long gone

by PETER TEMIN

ARE TEACHERS UNDERPAID? As an economist, I find it difficult to question market outcomes. Goods and services generally cost the market price. Only in the event of some kind of "market failure" do we say that goods or services are either over- or underpriced—meaning that they cost more or less than the price would have been in a fluid, competitive market.

At first glance, it is hard to see how the market for teachers could fail. True, most teachers' salaries are set by governments in a noncompetitive environment. But candidates choose freely whether to become teachers, in full knowledge of what salaries they will receive. In this sense, the people who choose to become teachers are paid a salary commensurate with their skills, preferences, and working conditions. These teachers are not underpaid relative to what they could earn in other occupations.

But what if we wanted to draw a higher-quality pool of candidates to teaching? Are we paying too low a price to accomplish that? I argue that the market for teaching has failed—in the sense that we are paying low salaries for low-quality teachers when we would prefer high-quality teachers. This is the result of two main flaws in the market: the difficulty of identifying who will be a good teacher and the reliance on an obsolete conception of the pool of potential teachers.

Squeezing the Lemons

Nobel Laureate George Akerlof's "lemons model" of market failure helps to explain why schools may not be willing to pay the market price for good teachers. Consider the market for used cars. Buyers want to purchase good cars, but they wonder if only "lemons" are put up for sale. And they have no easy way of evaluating whether a particular car is a lemon. They can look under the hood, take it for a test drive, even ask for the

alary on the
988 average
tartin or teach
usine stry.
19,60 ng
alarie
eache
or ch
21,900 for humanitie
majors. The large sa
%

owner's service records, but none of those will guarantee a vehicle's integrity. The buyer's concern for quality without a corresponding ability to evaluate that quality can cause markets to fail. Buyers may not be willing to pay the price necessary to draw high-quality cars into the market, thus ensuring that the proportion of lemons in the market will be higher than if the market were functioning well.

Likewise, schools are deeply concerned with the quality of their teachers, but quality is difficult to discern when hiring a new teacher. Studies have found that graduating from a good college and achieving high scores on tests of verbal aptitude are reasonable, though highly imperfect, indicators of teacher quality. In addition, licensing and certification rules attempt to ensure that teachers possess a certain level of skills. Nevertheless, many of the attributes that make for a good teacher are outside the bounds of a regime for testing or licensing teachers. A high-quality teacher is one who can energize and motivate students in addition to imparting information—qualities that are hard to recognize at the hiring stage. Teachers themselves may know how good they are, but the principals and school boards who hire them have far less information, as in the lemons model. They might desire a higher level of quality, but they are reluctant to pay the salaries necessary to obtain it because of the difficulty of choosing quality teachers.

The rise of a generation of powerful female politicians, including U.S. Environmental Protection Agency administrator Christine Todd Whitman, pictured here, is deeply symbolic of the dramatic change in women's opportunities during the past century.

PHOTOGRAPH FROM AP WORLD WIDE

Teachers are thus underpaid in the sense that we are paying low salaries for low-quality teachers. If we wanted, we could reach a different point in the market, where we would pay high salaries for high-quality teachers. As with used cars, a small salary increase would not change the quality of aspiring teachers; only a dramatic increase would attract a different pool of candidates and prove sustainable. It has been hard to make this kind of radical change because of historical patterns in the workforce that once allowed schools to educate on the cheap.

Women's Work

Women once considered teaching a highly attractive profession because their opportunities were tightly circumscribed. Despite the low wages, teaching was a far better line of work than slaving away in a textile mill. However, the past half-century opened a vast new world of opportunities to educated women. The nation's failure to accommodate these recent changes has kept teachers' salaries artificially low.

In Massachusetts the proportion of teachers who were women rose from half to more than 85 percent during the 19th century. By the 1920s almost all elementary-school teachers in the United States were women. They never were a large proportion of employed people in New England or even of employed women, but on average teachers taught longer than women worked in other occupations. The pool of female teachers comprised two quite different groups. Those who got married after starting out as teachers taught for about four years on average. But a large proportion of teachers did not marry. Thirty percent of teachers remained single, compared with 15 percent of New England women, after the Civil War. These single teachers taught an average of 12 years, raising the average tenure of teachers.

At the time, teaching was an unusually attractive career for women, compared with the alternatives. Almost two-thirds of Massachusetts women who earned wages in 1837 worked in their homes, producing goods such as palm-leaf hats, straw bonnets, and shoes and boots. The next most important employers of Massachusetts women in the 1830s were the cotton mills. As the century progressed, domestic service replaced braiding hats, but roughly half of employed women continued to work in textiles and the needle trades. This pattern began to erode during the early 20th century. As store-bought goods and services increasingly replaced those made in homes, women were increasingly demanded for work outside the home. Firms began to hire women to perform the market equivalent of homemaking: cleaning, cooking, and caring for the sick. The increasing need for clerical workers and the invention of the typewriter also afforded better jobs for women.

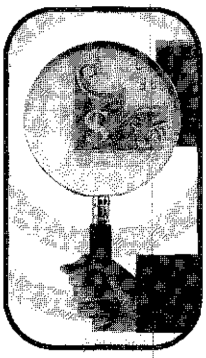
The participation of women in the labor force grew rapidly during the 20th century, from less than 20 percent of women in 1890 to well over 50 percent today. But women continued to have a narrow range of job choices for the first two-thirds of the 20th century. While the choices were not as restricted as in the 19th century, it should not be surprising that as late as 1960 many jobs held by women were still in cotton goods, clothing, and boots and shoes.

Only in the current generation have opportunities for women opened up. Previously, school boards hired female teachers in a context of little competition. Even in the mid-20th century, with opportunities still limited, teaching remained a good job, one that was considerably more interesting and probably more attractive than the alternatives. Schools didn't need to pay high wages in order to attract high-caliber individuals. High-quality women went into teaching anyway.

Then the market for teachers changed. The opportunities for women have expanded considerably, and their ability to get professional education has increased. In this new world, the brightest women go toward the best jobs. These jobs increasingly are not in teaching. In this kind of market, a higher salary is necessary to compete for higher-quality candidates, who now have other interesting career choices.

Harvard professor of education Richard Murnane argued a decade ago that there were many reasons why good candidates did not go into teaching, and salary was high on the list. In 1988, when the average starting salary for teachers was $19,600, business and industry offered average starting salaries of $28,900 for physics majors, $25,900 for chemistry majors, and $21,000 for humanities majors. The large salary gap for science majors surely discouraged them from entering teaching. Dedicated individuals and poor students who did not anticipate doing well in business still went into teaching, but low salaries discouraged many bright and vigorous students.

Using data on teachers' test scores from North Carolina, Murnane found that teachers with high test scores left teaching sooner than those who did less well. This is not surprising; we may presume that high-scoring teachers had access to better jobs than low scorers. Murnane also found that teachers who were paid more stayed in teaching longer. Taken together, these results showed that low salaries not only discouraged good students from becoming teachers, but they also discouraged good students from remaining teachers. The results of recruitment and retention are consistent; low salaries were associated with low-quality teachers.



In the mid-20th century, schools didn't need to pay high wages in order to attract high-caliber individuals. High-quality women went into teaching anyway.

New York University economist Frederick Flyer and the late University of Chicago economist Sherwin Rosen standardized teacher salaries for the demographic composition of teachers. They found that what they called the true real wage of teachers has declined relative to wages for other college graduates and has declined even more relative to female college graduates, especially for teachers in elementary schools. Most of this decline took place during the 1970s, when school enrollments fell, but relative wages have not recovered as enrollments have risen more recently.

To see what happened in the 1990s, I examined the relationship between average teacher salaries and women aged 35 to 44 with some graduate education. I compared teachers with women who had some graduate education to capture the comparison between teachers and women who have gone on to get degrees in subjects other than education (see the top line in Figure 1). Highly educated women who were not teachers earned more than teachers throughout the past 20 years. The wage premium stayed under 20 percent until around 1987, and it has risen steadily and strongly since then. The wage premium for women who have some graduate education and are not teachers is now 40 percent. This trend shows that teachers' salaries are falling behind those of highly educated women in other fields at a rapid rate.

Now, many high-school teachers have master's degrees, but most elementary teachers do not. Therefore, this comparison may not be accurate for teachers of lower grades; I may have compared them with women who have more education than they do. I cannot distinguish teachers by grade, but I can compare teachers' salaries with women who have only college degrees. This comparison is also shown in the bottom line in Figure 1. The trends are very much the same, although the divergence of wages is less strong in the bottom trend line than in the top line. Women with just a college degree earn an average of only about 10 percent more than teachers. Relative wages for elementary teachers fell in the 1970s and declined again in the 1990s.

Quality Control

The earnings comparisons tell only part of the story. The career choices open to women have expanded greatly in the past generation. This, coupled with the appearance of women with extensive experience in the workforce, has resulted in rising earnings for professional women. In addition, it has opened up jobs that are more interesting and challenging, careers that are more fulfilling. If there were a way to adjust for the expanding choices for educated women, the gap between teachers' earnings and alternative "earnings" would be larger than those shown here.

There also may be a sample-selection problem in the college-only comparison. If the brightest of women seek education after college, then the control group in the bottom trend line of Figure 1—women with only four years of postsecondary education—is self-selected to be of lower quality. The comparison in the figure reveals that teachers earn less than even those women who choose not to pursue their education beyond an undergraduate degree.

Finding themselves with lower-quality teachers, school districts have imposed work standards on teachers to make sure they are doing their jobs. These restrictions on teachers' creativity have made teaching an even less desirable job than is indicated by the salary. This is revealed most vividly in the relations between teachers and administrators. The latter feel obliged to keep tight control over teachers, and they set tasks for them to maintain minimum standards. Lesson plans and other reports increase the workload of teachers and force them into a uniform pattern. Of course, these work rules also discourage the most adventurous and creative teachers.

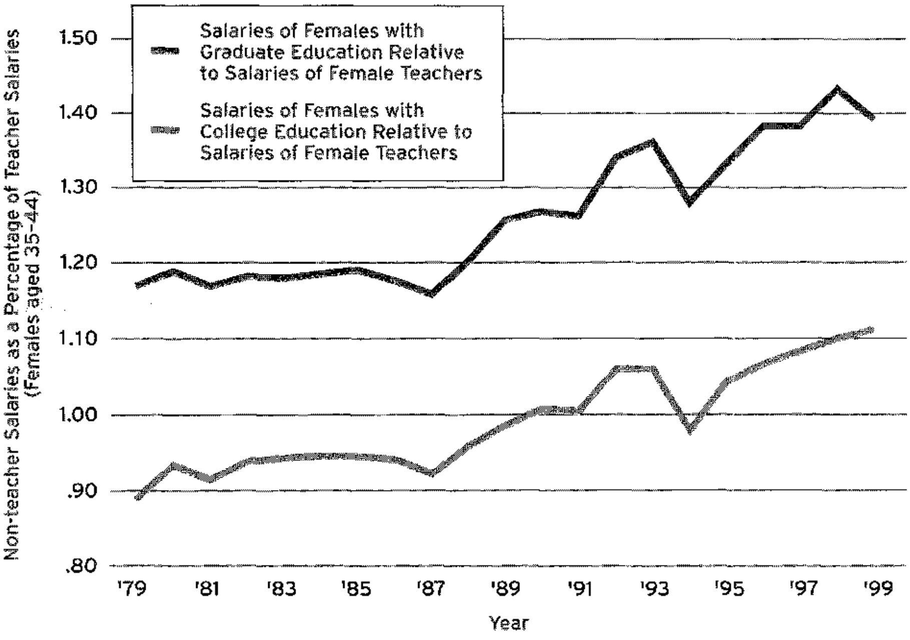
In some urban schools, principals feel free to discipline teachers in front of their students. This process humiliates the teachers and undermines their authority. Teachers are publicly criticized for not keeping their lines straight, for not being in "homeroom circle" properly, and for other minor infractions of classroom neatness. A Boston middle-school teacher told me that her principal yelled at her for forsaking her lunch duty place briefly to get food for a student whose lunch she had had to confiscate earlier. This is the kind of discipline that is designed to discourage all but the most dedicated of creative teachers.

Standardized tests, the cornerstone of the federal No Child Left Behind Act, are another way of monitoring teacher performance. The new purpose of the tests is to evaluate the schools and the teachers within it. Tests previously were used to evaluate students, but annual testing was unnecessary for this purpose. The new policies are meant to keep a tight rein on teachers rather than students.

Tests, like excessive paperwork and humiliating disciplinary actions, make life miserable for teachers. They force teachers to teach to the tests, reducing the scope of their teaching. They direct attention to skills that can be measured on annual objective tests and away from learning that is more subtle or long-term. For example, we charge the public schools with promoting civic responsibility. It is not clear how one would test this by multiple choice or if one would look for steady year-to-year increases in this set of attitudes. Less tangible and more creative parts of teaching are discounted relative to the basic mechanics of teaching. If low salaries do not drive creative teachers away, excessive discipline and annual tests should do so.

Pulling Away (Figure 1)

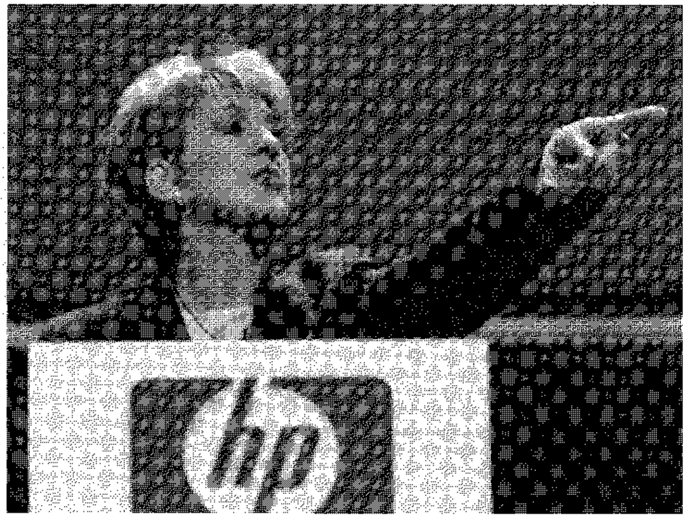
Women with graduate education in other professions now earn 40 percent more than female teachers, reflecting the expansion of employment opportunities for women during the past few decades. Even women with just a college education now earn about 10 percent more than female teachers.



SOURCE: Current Publication Survey

I do not mean to throw the baby out with the bathwater. Any dramatic increase in salaries will have to be accompanied by an increase in accountability. Inevitably, teachers' performance will have to be measured by some sort of tests. But the current crop of tests appears to be aimed only at evaluating minimum competence. Such tests tend to stultify the most creative teachers even as they, at least in theory, help the worst students.

Principals also may be more a part of the problem than the solution. They are recruited from the pool of teachers, and their quality has also suffered a decline. In the past, the demands on principals—to preserve order and maintain the school building—were consistent with their abilities. But if we wish to move to a point where good teachers will be rewarded for excellence and all teachers will be held accountable for their performance, it is far from clear that the current stock of principals will be adequate. To make tests an effective device for monitoring and rewarding quality, school administration as well as the tests themselves will have to



Fifty years ago, the only attractive option for a supremely talented woman like Carly Fiorina might have been to become a teacher. Today she is the CEO of Hewlett-Packard.

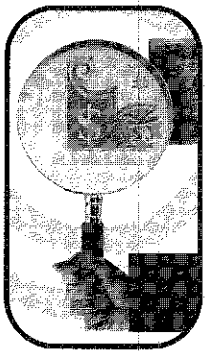
PHOTOGRAPH FROM AP WORLD WIDE

be improved almost beyond recognition.

Higher Aspirations

For most of U.S. history, the nation has been able to operate schools at low cost by exploiting the trapped labor force of educated women who had few other opportunities. Now schools must compete with other mentally and financially rewarding occupations as they recruit teachers. Surely it is immoral for us to shortchange our schools because we cannot bring ourselves to pay enough to attract good teachers.

The underpayment of teachers is likely to have important implications for the future of the American economy. The economy's success during the 20th century was due to many factors, of which the broad sweep of American education was an important one. The United States led the world in spreading education to the masses. The American economy grew consistently during the 20th century because we constantly upgraded the quantity and quality of our human capital. The United States has been successful both politically and economically because people have been able to transcend their origins and reach heights undreamed of in earlier times and other places. Education has been one of the primary paths through which extraordinary individuals have found an outlet for their gifts. We should try our best to preserve this commendable quality of American life.



Education reforms will be hampered by low teacher quality as long as the salaries of teachers do not rise to a competitive level.

Further education reforms will be hampered by low teacher quality as long as the salaries of teachers do not rise to a competitive level—the level that would attract many high-quality applicants. The current reforms of school administration and evaluation take the quality of teachers as given; they simply rearrange the existing educational assets, and as a result have little or no effect. Only when we break out of the current mix of pay and quality will education in the United States show a marked improvement.

Raising teacher salaries across the board would be very expensive. It also might take decades for higher pay to attract enough bright new teachers to make a big difference in the composition of the teaching staff. In the meantime, today's teachers would be paid more, with little likelihood that they would perform any better than they already do.

A more practical plan of action might be to couple increases in teacher salaries with tougher accountability. In an ideal scenario, only those teachers who were effective or who agreed to work harder to attain this standard would be retained. The pay hike then would not be as expensive, and if better teachers could handle larger classes, expenditures on teachers might not increase much at all.

How would we know which teachers are effective? Tests are imperfect at this stage, and they tend to measure achievement only on basic skills. It is easier to discover whether Mary can add than it is to find out if John is thoughtful and learns from his mistakes. Leaving the judgment up to principals is equally problematic, since they are part of the problem. While accountability is good, the nation's ability to implement it is imperfect. We need to find either a way to anticipate whether teachers will be effective or an acceptable method of evaluating their performance after the fact before we can implement a national program.

This all calls for experimentation. If teachers and their unions can agree, perhaps policymakers can institute bundles of pay raises and accountability in different forms in a variety of school districts. As the results of these pilot programs became known, they could be the basis for a more comprehensive call to action. Evolution may be more attainable and perhaps even more desirable than revolution.

–Peter Temin is a professor of economics at the Massachusetts Institute of Technology. This essay is adapted from an article that appeared in the Eastern Economic Journal.

comparable WORTH

Salary data fail to account for the shorter workday and work year in teaching. Once adjusted, teacher salaries look about right

by RICHARD VEDDER

IT IS AN ARTICLE OF FAITH AMONG MANY SUPPORTERS OF public education that teachers are underpaid. As Gayla Hudson, a former National Education Association official, once put it, "Until you start paying teachers at the level that other professions receive, recruiting will be a problem." Some respected academics have argued similarly. For instance, Peter Temin of the Massachusetts Institute of Technology (see "Low Pay, Low Quality" on page 8) claims that the opening of new job opportunities for women has drained the most promising candidates from the pool of potential teachers. In this view, higher pay is necessary to lure high-quality applicants away from more lucrative alternatives.

Most new teaching positions typically attract numerous applicants, especially in elementary schools, in subjects like English and social studies, and in the suburbs, where school districts frequently draw 100 applications

UNITED STATES OF AMERICA
THIS NOTE IS LEGAL TENDER FOR ALL DEBTS, PUBLIC AND PRIVATE
L 08352569 F
12
L
$
WASHINGTON, D.C.
12
L
L 08352569 F
12
12
%
8 percent

for a single slot. But the mere ability to fill positions is a poor measure of whether pay is adequate. The question is whether salaries are high enough to draw applicants of the caliber parents and policymakers desire. Furthermore, the benefits of higher salaries must be weighed against the benefits of other reforms that the money could be put toward. For instance, say a school district could raise its average test scores by one point by spending an extra $1 million on teachers' salaries. But what if the school district could induce a two-point increase in scores by spending the same amount on nonprofessionals to tutor struggling students? Or what if reforms outside education, such as better prenatal care, incentive-driven changes to public assistance programs, or job training, provided more bang for the buck in raising student achievement? In either case, public funds would be better spent on reforms other than increasing teacher salaries.

> Teachers earn more per hour than architects, civil engineers, mechanical engineers, statisticians, biological and life scientists, registered nurses, university-level foreign-language teachers, and editors and reporters.

The Evidence on Teacher Pay

A substantial body of evidence implies that teachers are *not* underpaid relative to other professionals. Using data on household median earnings from the U.S. Department of Labor, I compared teachers with seven other professional occupations: accountants, biological and life scientists, registered nurses, social workers, lawyers and judges, artists, and editors and reporters. Weekly pay for teachers in 2001 was about the same (within 10 percent) as for accountants, biological and life scientists, registered nurses, and editors and reporters, while teachers earned significantly more than social workers and artists. Only lawyers and judges earned significantly more than teachers—as one would expect, given that the educational training to become a lawyer is longer and more demanding.

Teachers, moreover, enjoy longer vacations and work far fewer days per year than most professional workers. Consider data from the National Compensation Survey of the Bureau of Labor Statistics, which computes *hourly* earnings per worker. The average hourly wage for all workers in the category "professional specialty" was $27.49 in 2000. Meanwhile, elementary-school teachers earned $28.79 per hour; secondary-school teachers earned $29.14 per hour; and special-education teachers earned $29.97 per hour. The average earnings for all three categories of teachers exceeded the average for all professional workers. Indeed, the average hourly wage for teachers even topped that of the highest-paid major category of workers, those whose jobs are described as "executive, administrative, and managerial." Teachers earned more per hour than architects, civil engineers, mechanical engineers, statisticians, biological and life scientists, atmospheric and space scientists, registered nurses, physical therapists, *university*-level foreign-language teachers, librarians, technical writers, musicians, artists, and editors and reporters. Note that a majority of these occupations requires as much or even more educational training as does K–12 teaching.

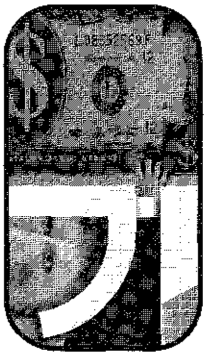
Government data on wages and salaries also exclude fringe benefits. Typically, teachers' retirement and health insurance benefits are more generous than the average professional's, particularly those who work in the private sector. Federal data suggest that, on average, teachers receive a package of benefits valued at more than 26 percent of their salaries. By contrast, the average for "all domestic industries" is about 19 percent; for private industries it is even less, below 17 percent. Take health insurance. Federal data suggest that about one-half of teachers pay nothing for single coverage (the employer pays everything), whereas the proportion of private "professional and technical" workers who pay nothing is only one-fourth. If direct hourly compensation averages perhaps 5 to 8 percent more for teachers than for all professional workers, and fringe benefits are perhaps 5 percent more, all told, teachers' average hourly compensation plus benefits exceeds the average for all professional workers by roughly 10 to 15 percent. In addition, teachers experience more job security, rarely suffering layoffs or firings. An architect's income falls with recession-induced declines in construction spending, while the economy is currently littered with unemployed computer programmers, whose jobs disappeared when the dotcom bubble popped. Most risk-averse persons would gladly accept some decline in average annual income in order to avoid unexpected adverse shocks to their standard of living. Thus, in a "risk-adjusted" sense, teacher pay is often substantially higher than that of comparable occupations.

Shortages

Of course, these are simple averages. It is possible that the pay for some kinds of teachers is relatively low. Others may receive salaries higher than they would earn in other professions. In my view, three groups of public school teachers can make a good claim for *not* being typically paid above-market compensation: unusually effective teachers, teachers in some geographic areas, and teachers with specialties that are in great demand.

In almost all private-sector jobs, pay is proportional to performance—to the estimated revenue that a worker's efforts bring to the firm. This theory of how wages are determined lies at the core of standard microeconomic theory and has been verified empirically in literally hundreds of studies. Workers are paid on the basis of *outcomes*. In public school teaching, by contrast, teachers are usually paid on the basis of *inputs*—the amount of schooling and experience that they possess. Qualitative differences in teaching ability are almost always ignored. Thus it is possible that teachers of average quality may be "overpaid" in the sense that they are earning more than workers of comparable training and productivity in other professions. Likewise, highly effective teachers might be "underpaid," earning less than equally skilled workers in other professions.

Unusually effective teachers, those in certain geographic areas, and those with specialties in great demand are typically not paid above-market salaries.



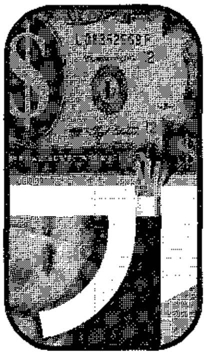
Critics of "merit pay" argue that it is impossible to evaluate a teacher's ability with any precision. Thus, allowing administrators to exercise subjective judgment could introduce major inequities in pay, since they will give merit raises to their friends and punish their enemies. Yet merit pay is commonplace in universities, and professors' teaching *is* evaluated. In a high-quality university, salaries among professors with the same formal education and years of experience might differ by as much as 100 percent or more, and differentials of 50 percent are not uncommon within the same academic department. Even in schools of more modest repute, top-flight faculty members may make 50 percent more than those perceived as less productive. In K–12 schools, a whole variety of indicators, such as test scores, enrollment figures, and college placement statistics, could be used to evaluate teachers, yet this is seldom done. By the way, this problem has a long pedigree. Writing in 1776 in the *Wealth of Nations,* Adam Smith complained, "The usual reward of the eminent teacher bears no proportion to that of the lawyer or physician." He attributed this to the move away from market-based funding of education, noting that in ancient Greece, where all education was private and financed by tuition payments, great teachers were often extremely well paid relative to their less-talented counterparts.

Average teacher salaries vary enormously in most states by area, tending to be higher in suburban and urban districts and much lower in rural areas. In my state of Ohio, for example, the average teacher salary in the district with the highest pay is about double that of the district with the lowest pay, a differential that cannot be explained entirely (or even mainly) by differences in teachers' characteristics. Applicants are often abundant in high-achieving, high-salary suburbs, while school principals and human-resource managers in poor, relatively isolated rural school districts may face genuine teacher shortages. If their salaries were higher, they might be able to attract teachers from places that are more desirable to live.

Interstate variations in average salary are likewise substantial. In 2000 the average public school teacher salary reported by the National Education Association was less than $30,000 in North and South Dakota, but more than $50,000 in New York. Yet these differentials are often justified by local labor-market conditions. The average teacher made $51,800 in Connecticut, but only $35,000 in West Virginia. Yet, in a relative sense, the West Virginia teacher earned a higher salary. Teachers in West Virginia earn 30 percent more than the average workers in West Virginia, while teachers in Connecticut earn only 11 percent more than the average worker in their state. Generalizations about geographic inequalities in teacher salaries typically ignore these profound differences in labor markets.

Shortages have developed in fields like mathematics, the sciences, and special education in large part because school districts make no distinctions among teachers. According to the rigid salary schedules typically negotiated by the teacher union, those who are in high-demand specialties must earn the same as teachers in fields where there is a surplus of applicants. Thus schools that have no trouble finding English, history, or elementary teachers may have to scramble to hire special-education or science teachers. At a typical university these days, a new assistant professor in my field, economics, will earn 50 to 100 percent more than a new assistant professor of English or history. By contrast, in a typical public school district there is *no* differential, and new English teachers may be paid perhaps $5,000 or more than their "reservation wage," or the lowest wage for which they will work. Meanwhile, the district may have trouble finding any special-education teachers, indicating that the going rate is below the reservation wage of competent potential employees.

In short, even where there are documented instances of teacher shortages, the problem is seldom that teacher salaries overall are too low, but rather that inefficient practices ignore the realities of the market. Without changing average teacher salaries a penny, most districts could ease their recruitment problems by paying some teachers (good ones, those in high-demand specialties) more than others (less-competent ones in low-demand specialties). The problem is not the *level* of pay, but rather its *distribution*.



The unions' persistent quest to raise teacher salaries looks more like a ploy to redistribute income from the general taxpayer to teachers.

Public versus Private

The most compelling evidence that teachers are not on average underpaid comes from the world of private schools. The last comprehensive analysis, performed during the mid-1990s, indicated that average private school salaries were slightly less than 60 percent of average salaries in the public schools (though the gap between public and private schools narrows substantially when religious private schools are excluded from the analysis. See Michael Podgursky's article, "Fringe Benefits," on p. 71 of this issue). There is no reason to believe that this differential has changed dramatically in the past few years. In other words, a public school teacher might make more than $40,000 in annual salary, while a private school teacher with similar levels of experience and education would earn $25,000.

Are public school teachers any more effective than those in private schools? On the contrary. In general, children in private schools perform better on standardized tests, even after adjusting the data to reflect differences in other factors that are relevant to learning, such as family income. The public school crowd retorts that public school teachers are certified by their state, indicating that they have taken the requisite number of courses in pedagogy, while most private school teachers, especially at the secondary-school level, lack certification. But there is little to no evidence that certification improves the quality of instruction. The absence of certification requirements is actually a virtue of the private school market, since it gives private schools access to a larger pool of qualified applicants.

Teacher salaries are set by local school boards, which compete with one another for talent. This suggests that salaries should be the going rate. So how can private schools get high-quality teachers for 40 percent less pay than the public schools can? For one thing, it is important to acknowledge that many religiously motivated teachers will sacrifice higher pay to work in a private school whose mission comports with their core beliefs. Secular private schools tend to pay more than religious schools. However, at least part of the answer lies with the powerful influence of the teacher unions. Through negotiations, political pressure, and the threat of strikes, the unions have been able to secure wages that are higher than would be dictated solely by market considerations. From this perspective, the unions' persistent quest to raise teacher salaries, carried on under the banner of improving education, looks more like a ploy to redistribute income from one group of adults—the general taxpayer—to another—teachers.

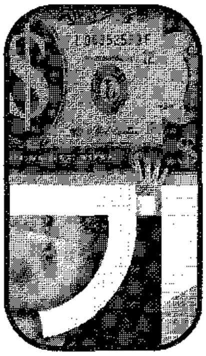
It is probably no accident that the differential between teachers' pay and the pay of workers in general accelerated during the 1960s, the decade when unionization among teachers took hold. Some scholarly studies suggest that unionization has raised teacher salaries (and per-pupil costs) by perhaps 10 percent. If accurate, the average pay of the 3 million public school teachers may be $5,000 or so higher annually because of unionization, or $15 billion for the entire teacher corps.

It may be that the premium paid to public school teachers reflects both the influence of their unions and the more difficult environments they work in. Public school teachers enjoy far less freedom than their private school counterparts; their actions are severely circumscribed by laws, court decisions, and administrative dictates from remote bureaucracies. Unruly students cannot be expelled or adequately disciplined. Teachers often cannot use the instructional materials that they prefer. They have little or no say in the hiring of colleagues, the establishment of curricular goals, or the disciplining of students. There is a sense of frustration, a lack of trust and of teacher empowerment. Teachers in private schools work in an environment that is comparatively less bureaucratic. They enjoy school communities where collegiality is high, goals are well stated, and success is rewarded.

In other words, the salary differential between public and private school teachers may be explained partially by what is often called "combat pay." There is some evidence to support this. In surveys of private and public school teachers taken during the 1990s, public school teachers consistently showed less satisfaction with their work conditions than their private school counterparts. For example, on a 1993 survey, 21 percent of public school teachers did not agree with the statement, "The school administration's behavior toward the staff is supportive." Only 12 percent of private school teachers did not agree with the statement. Some 23 percent of public school teachers did not agree that "there is a great deal of cooperative effort among staff," compared with only 10 percent in private schools. Nearly half of the public school teachers thought that they did not receive a great deal of parental support, compared with only about 12 percent of their private school counterparts. "Student disrespect for teachers" was a serious problem for 19 percent of public school teachers, compared with only 3 percent of private school teachers.

Yet when asked about their salaries in the 1993 survey, teachers in both private and public schools reacted similarly. Despite earning far less, teachers in private schools were only slightly less satisfied with their salary than their public school counterparts. This is highly consistent with the idea that some of the premium paid to public school teachers is due to less desirable working conditions.

Currently, there is no salary premium associated with taking on larger classes, so teacher unions fight for smaller classes.



Teacher Pay and Education Policy

I suspect that the failure to link teacher pay to individual teachers' productivity has driven the education policy debate in unexpected ways. Consider the move to reduce class size. Teachers and their unions have been major proponents of this "reform." But would that be the case if teacher pay were tied to productivity? I doubt it.

Suppose a high school has two teachers of junior and senior classes in history, each with 100 students in five classes averaging 20 students each. Suppose both teachers make $50,000 a year, including fringe benefits. Then suppose that the less effective of the two teachers retires. Typically, the school would hire a replacement, costing perhaps $40,000 a year, because he has less experience. Now, suppose that instead the school simply asks the more effective history teacher to take on the retiring teacher's load, thereby doubling his class sizes to 40 students. The teacher would typically howl in protest. But what if the school district offered the teacher $15,000 in additional pay and fringe benefits, plus a part-time teacher's aide paid $10,000 a year to assist in grading and record keeping? The teacher would probably accept the assignment. The district would save $15,000 and probably maintain the level of instruction. The excellent teacher is likely to be better at teaching 40 students per class than an unknown, less-experienced teacher would be at teaching 20 students per class. Currently, there is no premium associated with taking on larger classes, so teacher unions fight for smaller classes. But sharing some of the cost savings of larger classes with teachers may be one way of breaking down teachers' resistance to changes in the status quo. The resistance to larger classes is not just among teachers. Private schools typically have relatively small class sizes—and affluent parents seem to want them even if it adds significantly to tuition costs.

Market-based reforms would also force changes in the system of teacher compensation. Earlier I argued that public school teachers receive a pay premium for two reasons: the influence of their unions and the combat pay necessary to compensate them for their less than ideal working conditions. Market-based reforms would reduce the premiums associated with both aspects of the current public school system. Expanding the use of publicly funded vouchers and increasing the number of charter schools would reduce the power of the unions as more and more students moved to nonunionized private and charter schools. In turn, more and more teachers would move to private and charter schools, where their work environments would probably improve because of the stronger sense of school community that comes with having chosen a school. If taxpayers are paying a huge premium to public school teachers merely because they belong to powerful unions and work in unpleasant environments, then the costs of failing to reform the system are indeed substantial.

–Richard Vedder is a professor of economics at Ohio University and a senior fellow at the Independent Institute.

Façade of Excellence

Stuyvesant High, one of New York City's elite exam schools, enjoys an ultramodern building, influential alumni, and the city's brightest students. If only it could also escape ***the corrosive rules of the teachers' contract***

by SOL STERN

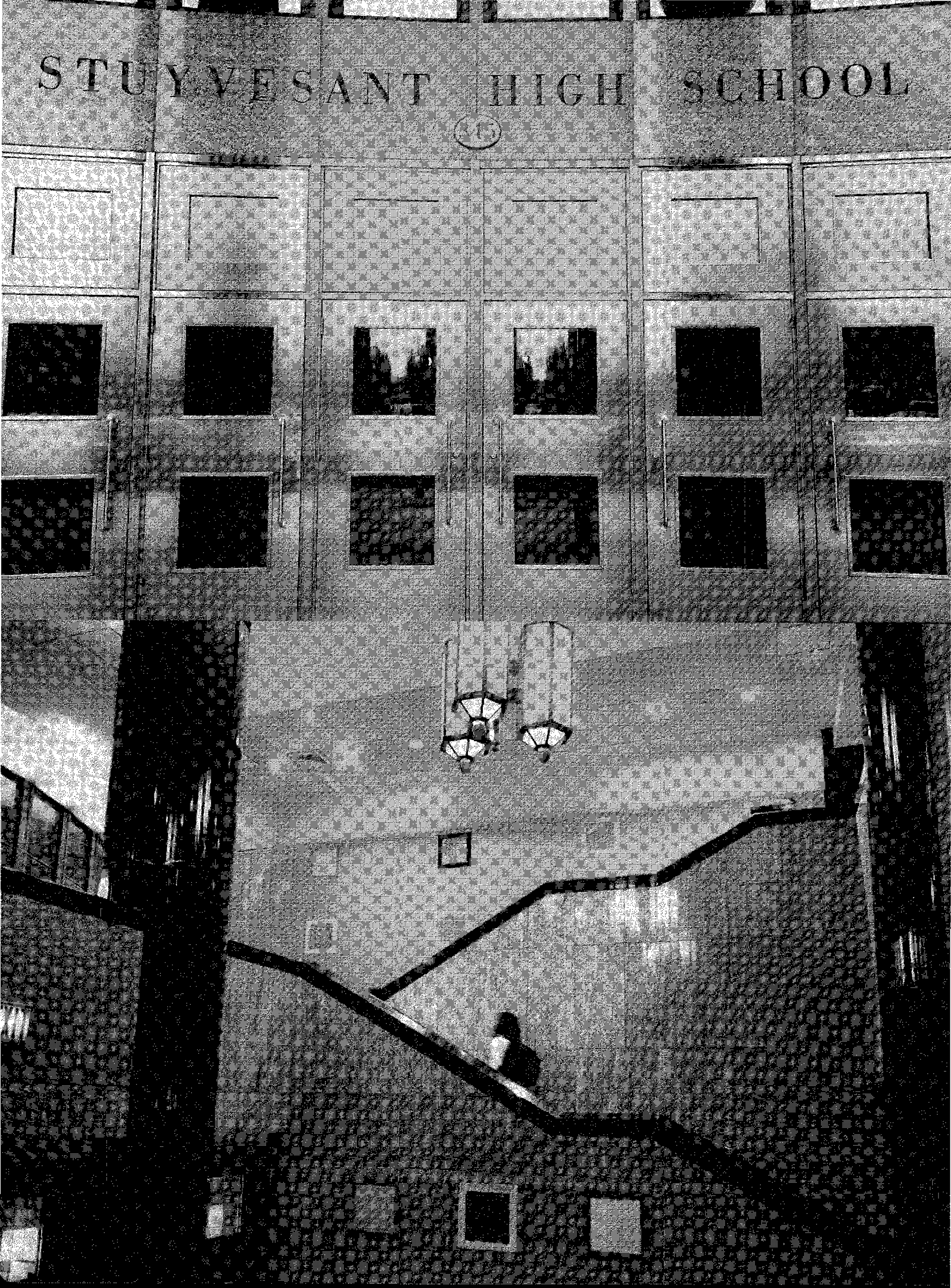
ON A BITTERLY COLD Saturday morning in December 1995, I wished my son Jonathan good luck and then watched him join a long line of students at the entrance to Stuyvesant High School in lower Manhattan. They were all waiting to take the admissions test for New York's three vaunted "exam schools."

The admissions process for the three schools is one of the last bastions of pure meritocracy in American education. Each year about 20,000 of the city's 90,000 8th graders report for the three-hour test of math and verbal aptitude. The grand prize, a seat at the world famous Stuyvesant, is offered to those with the top 800 scores. The next 600 are admitted to the Bronx High School of Science. Another 1,000 students get into Brooklyn Technical High School. There are no "legacy" or affirmative-action admissions. Political connections can't get your child in. Special consideration is given only to a small group of students who qualify as "economically disadvantaged" and come within a few points of the cut-off score. These students are given a summer-school course and allowed to take the test again. If they pass on the second try, they're in.

The three exam schools have been in existence for more than 70 years. During the first half of the 20th century, when the New York City public schools composed one of the most successful systems in the nation, they were considered the jewels in the crown. More recently they have been widely regarded as among the last exemplars remaining in a failing system.

The high-school options now available in the city are so limited that thousands of middle-class and working-class parents find themselves left out in the cold when their children fail to make the cutoff for the exam schools.

PHOTOGRAPHS BY PATRICK HARBRON


STUYVESANT HIGH SCHOOL

Some then move to the suburbs, while others endure heavy financial sacrifices to send their children to private or parochial schools. Attending a regular New York City public high school is usually the option of last resort.

Of course the situation could always be worse. Indeed, it would be much worse if John Lindsay, New York's iconic liberal mayor during the late 1960s and early 1970s, had gotten his way. Lindsay complained that the entrance exam for the three specialized high schools was "culturally biased" against black and Hispanic children and sought to have the high schools eliminated. He believed that it was a serious civil-rights violation for the city's elite public schools to use race-blind admissions standards, while his own children attended exclusive private schools where money, connections, and "culture" played a major role in the admissions process.

Iftimie Simion was soon teaching college-level math courses at Stuyvesant, yet his job was in danger because he had not accumulated enough credits to qualify for a teaching license.

PHOTOGRAPH COURTESY OF IFTIMIE SIMION

Fortunately, parents of students at Stuyvesant and the other exam schools were able to persuade two obscure state legislators from the Bronx to intervene. They sponsored a bill that later became Article 12 of the New York State Education Law, which reads: "Admissions to the Bronx High School of Science, Stuyvesant High School, and Brooklyn Technical High School shall be solely and exclusively by taking a competitive, objective, and scholastic achievement examination, which shall be open to each and every child in the city of New York."

The Best and the Brightest

Two decades later Stuyvesant High School received another unexpected gift: a brand new building. When I attended Stuyvesant back in the 1950s, our turn-of-the-century building on East 15th Street was so overcrowded that we went to school in shifts. By the 1980s, the building was falling apart. A committee of Stuyvesant parents and alumni then lobbied state and city officials for a new home. Remarkably, the city agreed to spend $150 million for an ultramodern building on a sliver of land next to the Hudson River in lower Manhattan.

The new Stuyvesant was designed by one of New York's leading architectural firms. On our initial tour, my wife and I couldn't believe that we were in a city high school. A pair of marble staircases sweeping up to the second floor greeted us as we entered the building. Two classic black columns soared to the top of the ten-story structure. The school boasts a competition swimming pool, three gymnasiums, a high-tech theater, and ten fully equipped science laboratories. The fifth-floor cafeteria has a balcony with a vista of New York harbor and the Statue of Liberty. Best of all, the building has central air-conditioning, a luxury among city schools.

The principal of Stuyvesant had no say in filling as

Each year, the parents association raises more than $300,000 to help pay for a smorgasbord of extracurricular activities. The school offers a 100-piece symphonic orchestra, a symphonic band, a Renaissance choir, and a number of chamber-music ensembles. The debate and math teams regularly finish first in national and regional competitions. Despite its reputation as a refuge for "nerds," Stuyvesant also fields more athletic teams than any school in the city, including a successful football program.

When my wife and I learned that our son Jonathan, and later our younger son, Dani, had been admitted to Stuyvesant, we felt as if we had won the state lottery. They would be surrounded by 3,000 of the city's brightest students, representing more than 50 different nationalities and ethnic groups. The students seemed to embody the spirit of an achievement-driven city and its succeeding waves of striving immigrants. Though many of them did not count English as their first language, the graduating seniors regularly attained average SAT scores of around 1400. Stuyvesant also frequently produces more National Merit Scholars and semifinalists for the prestigious Intel Science awards than any other school in the nation.

Yet all these accolades tend to mask the school's dirty little secret. While Stuyvesant admits only the city's finest students, through a process that is free of subjectivity and nepotism, its hiring decisions are far less merit-driven. Like every other public school in the city, Stuyvesant is plagued by bureaucratic regulations and corrosive work rules that favor seniority and paper credentials over a teacher's knowledge and skill in the classroom. For instance, according to the teachers' contract, half of Stuyvesant's teaching vacancies each year must be posted and set aside for teachers seeking transfers from other city schools. These set-asides are required to be filled solely on the basis of seniority. The fact that many of those selected lack the academic qualifications to teach to the level of Stuyvesant's students is irrelevant to the union and the school system. The underlying premise of the contract is that

Former Stuyvesant principal J. Cozzi Perullo (left) didn't like the seniority system and the work rules. And she said so in public. For this she made many enemies.

many as 50 percent of each year's teaching vacancies.

any teacher with a state license is fit to teach at Stuyvesant.

The Case of Iftimie Simion

My son Jonathan had an improbable streak of good luck with his freshman-year teachers. The most memorable was Iftimie Simion, a refugee from Romania, who was also in his first year at the school. Simion had taught at the university level for 13 years, coached the Romanian national math team, and published papers in professional journals. By the end of Simion's first year of teaching, Stuyvesant's administrators knew they had a rare gem. Simion was soon teaching college-level courses such as multivariate calculus and differential equations to the most advanced math students. He also coached the math team and supervised the math research program. Math department chairman Danny Jaye, a big baseball fan, alternately called him "my MVP" and "my indispensable franchise player."

In the public school league, however, talent doesn't count for much. As a recent immigrant, Simion hadn't yet accumulated enough education credits to qualify for a state teaching license. A citywide teaching shortage had caused the Board of Education to relax the hiring rules, enabling Stuyvesant's principal to hire unlicensed teachers like Simion. Such teachers, however, were vulnerable to being bumped from the staff if a senior teacher from within the system applied for their positions, which were considered vacant until they attained a permanent credential. One transfer applicant who came close to bumping Simion was a 71-year-old math teacher from Seward Park High School. Though he had logged 35 years in the system, this teacher had never taught at a level higher than first-year Regents math—which 90 percent of Stuyvesant's freshmen have taken in junior high school. On his first attempt to transfer, members of Stuyvesant's faculty were able to persuade him that he was not up to the task of teaching the accelerated Stuyvesant curriculum. At the last minute the elderly teacher declined the transfer that was his by right.

But that wasn't the end of the story. A year later he was again approved for a transfer to Stuyvesant. This time he decided not to let the administrators at Stuyvesant intimidate him. After all, didn't the union contract stipulate that he was entitled to transfer to any school he preferred? Now 72, and apparently beyond embarrassment, he joined the Stuyvesant math faculty. Fortunately, by then Simion had received his license, and his job was safe.

The students were the vulnerable ones now. The transfer teacher was in his third year of teaching at Stuyvesant when my younger son, Dani, was assigned to him. Dani described him as a decent, gentlemanly man who struggled to get things done in each 43-minute period. Whatever he did came straight from the textbook, and he avoided the more complicated problems or proofs. When I observed his classroom, it was obvious to me that he was completely out of his league. He worked at a snail's pace with kids who were ready to sprint. He used a crib sheet to steer him methodically through the steps of the solutions to problems he had placed on the board.

At least 16 percent of the total Stuyvesant staff was composed of teachers who received their jobs because of seniority or race and ethnicity rather than professional accomplishments.

With this math teacher the seniority transfer issue at Stuyvesant reached a level of comic absurdity. His lack of mastery was so glaring that he had to sit in on two lower-level math courses at the school. For an entire semester he went to classes, took the tests, and was given grades just like any other student.

Making Enemies

The issue of transfers became Stuyvesant principal J. (Jinx) Cozzi Perullo's bête noire—and ultimately led to her resignation. She had begun ruffling feathers from the moment she took the school's helm in 1994. She didn't like the seniority system and the work rules, believing they undermined excellence. And she said so in public. For this she made many enemies.

Perullo once described to me a conversation she had with the outgoing principal, Abraham Baumel, in which he told her that she now had the best principalship in the city. At first Perullo was flattered, until she realized what he meant: that Stuyvesant was a great place to be a principal because there was virtually no way to fail. The students guaranteed the school's academic success. No matter what the principal did, the students would still achieve average SAT scores of 1400 points, and 99.5 percent would go to college.

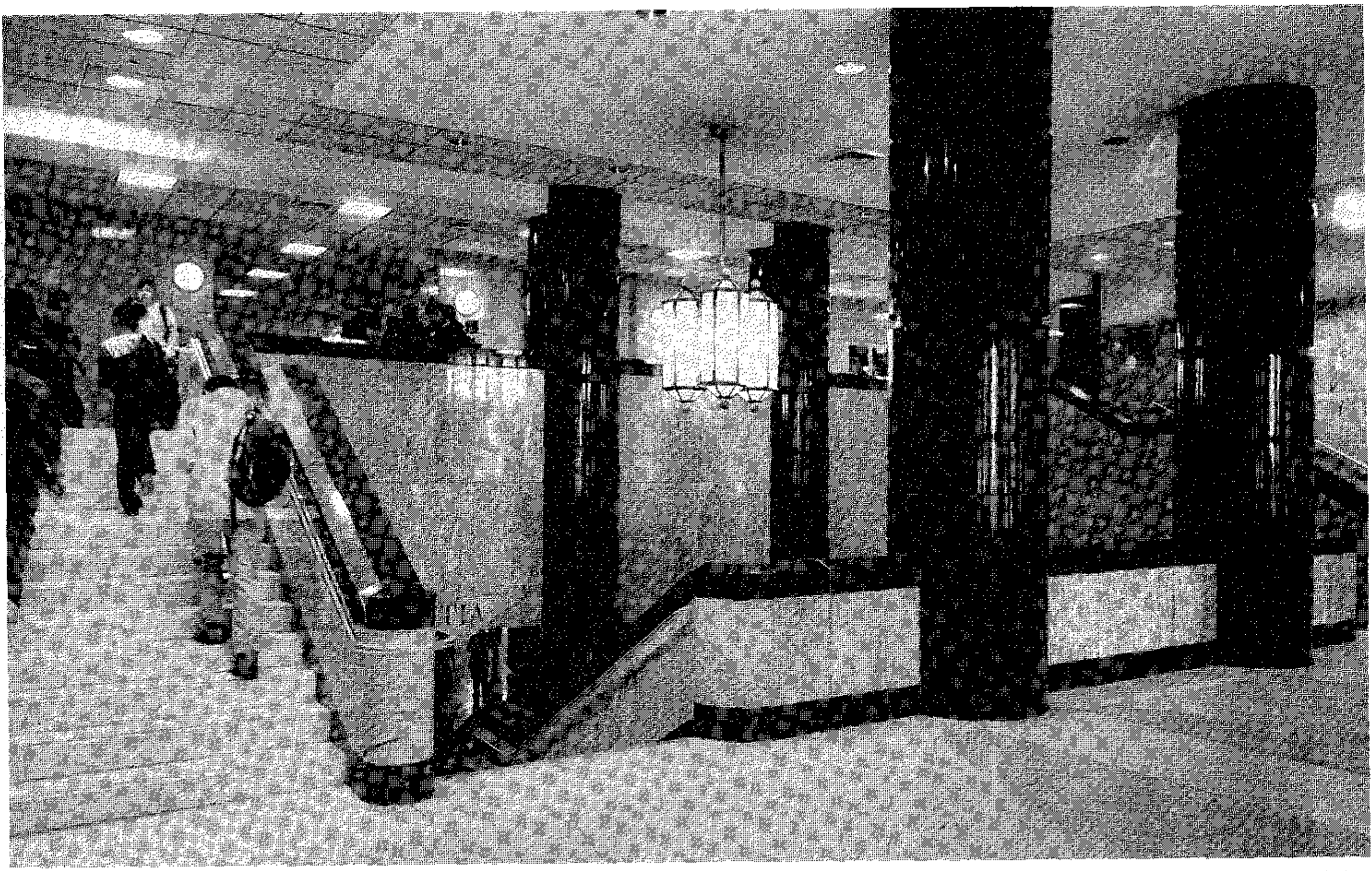
Perullo conceded that Baumel had a point. "There's never been a discipline issue here," she told me. "The kids always do their homework. In some ways Stuyvesant isn't a 'real' school. A teacher could fake it here for 35 years, because even when the teaching is inadequate the kids will find a way to do well on their tests."

Nevertheless, Perullo was determined to challenge the culture of the union contract. And by challenging the union's prerogatives she incurred the wrath of Dorothy Suecoff, a veteran biology teacher who served as the school's United Federation of Teachers (UFT) chapter leader. (Suecoff declined my requests for an interview.) As the enforcer of the union contract, she was in some respects as powerful as the principal. "Her role in the school is to defend the indefensible," said one of Stuyvesant's assistant principals.

The biggest weapon in Suecoff's armory was the elaborate grievance machinery available to the union if a principal should disregard any of the contractual work rules. Once Perullo made clear her intention to try to change some of Stuyvesant's old ways, Suecoff began filing grievances right and left. She objected to almost every directive coming out of the principal's office. When Perullo proposed experimenting with block scheduling (double periods for some classes), Suecoff countered with a grievance. When the student newspaper ran an article critical of the union or a particular teacher, a grievance was filed against Perullo on the grounds that the principal was supposed to protect teachers from student criticism.

At the heart of this increasingly bitter confrontation was Perullo's refusal to drop the issue of the seniority clause. The *New York Times* quoted her complaining that she had no say in filling as many as 50 percent of the teacher vacancies at Stuyvesant. Without the power to hire and fire teachers, she insisted, principals couldn't do their jobs.

Another part of the contract that troubled Perullo was the "Transfers to Further Integration." This granted transfer rights for minority teachers into any school that was deemed not to have a fully integrated teaching staff. It was affirmative action plus. Since there was a dearth of blacks and Hispanics in the profession, it meant that a school like Stuyvesant was forced to set aside its openings for any licensed black or Hispanic teacher who put in for the job. Perullo supported the goal of



The new Stuyvesant building, designed by one of New York's leading architectural firms, boasts a competition swimming pool, a high-tech theater, and ten fully equipped science laboratories. The fifth-floor cafeteria has a balcony with a vista of New York harbor and the Statue of Liberty.

bringing in more minority teachers, but she bridled at the fact that the process was corrupted by the seniority rule. Integration transfers frequently displaced younger, dedicated minority teachers at Stuyvesant who had been groomed by the principal, but had yet to earn tenure rights.

For obvious reasons, seniority and "integration" transfers are especially a problem at schools like Stuyvesant. After all, what veteran teacher wouldn't want to end his or her career in a luxury building among students who virtually teach themselves? According to Perullo's count, between 1995 and 2000 Stuyvesant took a total of 24 UFT transfers. In other words, at least 16 percent of the total Stuyvesant staff was composed of teachers who received their jobs because of seniority (or race and ethnicity) rather than professional accomplishments.

In the end, Perullo's endless battling with Suecoff took its toll. "On a personal level, I had never experienced anything like this in 28 years as a teacher and administrator before I arrived at Stuyvesant," she told me. "I was not accustomed to my person coming under attack because of my position as principal." The union leader's obstructionism did not go unchallenged, but when Suecoff won a re-election for union chapter leader that was essentially a referendum on her tactics, Perullo realized that Suecoff had enough support from the faculty to continue erecting roadblocks. Perullo soon announced her retirement after five years at the school. She minced no words in explaining that the union leadership made it impossible for her to fulfill her goals for the school.

Perullo's sudden resignation received a surprising amount of media coverage. In interviews she continued to press her contention that principals needed the power to hire and fire teachers to accomplish their mission. This was so abhorrent to Suecoff that she was still waging war against Perullo a year after Perullo's resignation. Suecoff introduced a resolution at the union executive committee that said Perullo (who still showed up at Stuyvesant from time to time because she was assisting her successor and had other unfinished business) should be barred from the building.

Reunion

As a journalist and frequent critic of the New York City school system, I was invited to participate in a televised forum on the public schools sponsored by the New York City Council. During a discussion on teacher quality, I cited my son's amazing math teacher, Iftimie Simion, as an example of how "uncertified" doesn't necessarily mean "unqualified." This had the serendipitous result of reuniting me with Mr. Simion and teaching me yet another lesson in the sheer stupidity of the New York City school system.

Simion had watched the show and contacted me, seeking advice. He was now in his fourth year at Stuyvesant. He had completed all the ridiculous education courses the system required, including credits in special education and "human relations." In return the state of New York had finally handed him a piece of paper that said he was qualified to teach high-school math, as he had already been doing brilliantly. But a Board of Education bureaucrat somewhere deep in the bowels of 110 Livingston Street had decided that it was too complicated to evaluate Mr. Simion's doctoral-level training in Romania (partly because some documents had been lost in his flight from his homeland). Thus Mr. Simion was denied the salary increment given to teachers who have attained a master's degree plus 30 credits.

He was also penalized by one of the most senseless clauses in the teachers' contract. It stipulated that a teacher coming into the system from any other school district (either public or private) could not be given salary credit for more than five years of previous teaching experience. The cap on experience had been put into the contract at a time when there was a surplus of teachers rather than a shortage. Even the UFT was now against the clause and proposed eliminating it (the policy has since been amended to credit teachers for up to eight years of experience).

All told, Simion was earning about $20,000 less than he should have been. His salary was also $35,000 less than that of the incompetent 74-year-old math teacher who had come close to replacing him. Meanwhile, two nearby suburban districts were offering Simion a higher salary scale and said they would recognize his previous years of experience in Romania. (Stuyvesant physics teacher Van Caplan had left for the suburbs a year earlier partly because the city refused to recognize his 20 years of teaching at West Point.) Although Mr. Simion loved teaching at Stuyvesant, he also had a family to support, and his son was about to enter college.

It occurred to me that if common sense didn't work at the Board of Education, plain embarrassment might. I called David Klassfeld, then Chancellor Harold Levy's top deputy and, coincidentally, a Stuyvesant parent. I knew that Klassfeld was aware of Simion's reputation, and Klassfeld knew what it would say about the Board of Education if Simion left the Stuyvesant math department because of the idiotic rules. Klassfeld also knew that I would write about the situation in whatever media outlet afforded me the space.

TRIBECA BRIDGE

Stuyvesant frequently produces more National Merit Scholars and semifinalists for the prestigious Intel Science awards than any other school in the nation.

PHOTOGRAPH FROM AP/WORLD WIDE

Shortly thereafter, Klassfeld called back to tell me that a partial solution had been found for Simion. A committee of outside experts had been convened at the Board to evaluate the coursework of teachers who had attended foreign universities. In fact the committee had already looked into Simion's case and ruled that he was entitled to the salary increment for education. He would even be paid the increment retroactively, dating back to his first year of teaching at Stuyvesant. The back payment of about $16,000 plus the extra $4,000 added to his base salary persuaded Mr. Simion to stay on at Stuyvesant for at least one more year. Two years later, however, he left for the suburbs, partly because he was fed up with a system that considered every math teacher with a license the same as every other teacher, no matter how much talent they brought to the classroom, and that never fully compensated him for his many years of high-level teaching experience.

Nevertheless, it was amazing to see how quickly the rules can change when there is the will to do so. I felt that I had made a small contribution to Stuyvesant's legacy of math and science excellence. I decided that I should do even more to undermine the oppressive rule of seniority and mediocrity at the school.

Obstruction

Stuyvesant had just recently put into place its School Leadership Team (SLT), a reform mandated by the state. This was supposed to be a vehicle for parents, teachers, and students to collaborate in school improvement. In theory, if the SLT reached a consensus, it could make major decisions about school practices that the principal was then required to carry out. The parents association elected the parental representatives, and I chose to participate as an observer with the intent of pressuring the elected representatives to raise the issue of seniority.

Still, I had few illusions about the SLT's potential as a reform tool. There was a big catch-22 in the enabling legislation: an SLT could not adopt any policy that conflicted with the union contract. The first SLT meeting I attended had the look of a United Nations Security Council meeting. The meeting opened with many delegates expressing frustration that almost nothing had been accomplished in their year of existence. This was attributed, much as it might have been at the U.N., to excessive procedural wrangling or to the power of one permanent member. Dorothy Suecoff had apparently played the role of the many Soviet ambassadors to the U.N. during the cold war, who just kept saying *nyet* and vetoed just about every resolution passed by the Council majority. The SLT would work through several meetings on an issue and arrive at a consensus, only to be told by Suecoff that the proposed change violated the teachers' union contract.

In one such instance the SLT had adopted a resolution calling for the principal to start an experiment with block scheduling. The idea was to have some teachers voluntarily teach some classes for two periods in a row for an hour and a half instead of one 43-minute period every day. (This was essentially the same proposal Jinx Perullo had made three years earlier.) A tremendous amount of time and effort went into drawing up the plan, including a visit by some delegates to Thomas Jefferson High School in Virginia, which uses block scheduling. But the proposal was killed at the end of this elaborate process when Suecoff filed a grievance with the superintendent of Manhattan High Schools. She claimed that the block scheduling would have meant that some teachers actually spent more than the contractual limit of teaching five periods in one day—even though this would be balanced by teaching fewer than the five periods on other days. Suecoff's grievance was upheld, even though the teachers in question would have *volunteered* to teach the block schedule.

> **"Every time the students try to bring up a program to improve education, we are told it violates the teachers' contract," said student Matt Kelley. "It seems like education improvement is itself a violation of the contract."**

To try to head off a similar fiasco in the future, one of the SLT subcommittees proposed a resolution requesting delegates to inform the SLT early on if they believed that a proposed initiative would "contravene a union contract or state education law." The resolution was clearly aimed at the UFT chapter chair, the only person who was invoking the union contract to kill proposals. Suecoff was actually on the subcommittee that had proposed the resolution, and everyone believed she had agreed to it.

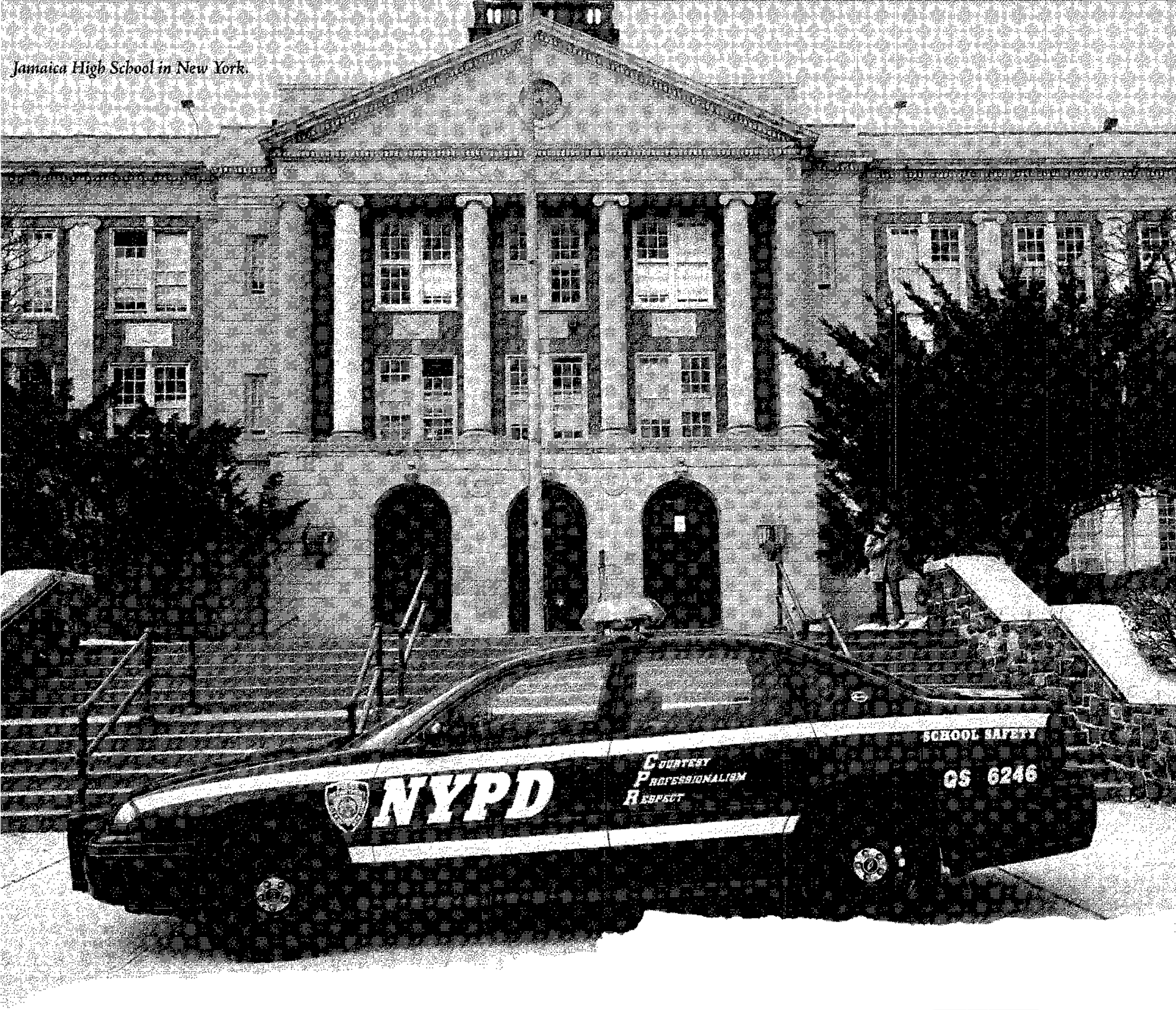
Suecoff sat expressionless as the resolution was read. She then made the following announcement: "I can't be bound by this resolution. I represent the staff. I have the right to file a grievance against this committee." Many of the delegates were shocked by this response. As one parental delegate put it, no one disputed the right of the UFT chair to file grievances. All that was asked of Suecoff was to let the delegates know when she thought they were headed into territory covered by the union contract. Suecoff wouldn't even offer her fellow delegates that crumb. She insisted that she wasn't going to provide the SLT with the benefit of her interpretation of the contract. "A grievance is a private matter," she added with emphasis.

This provided an enlightening introduction to the politics of school reform for the five student representatives on the SLT. The leader of the student delegation was Matt Kelley, a tall, handsome senior with longish blond hair who frequently mocked the principal's pretensions about Stuyvesant's being a great science and math school. "A lot of us came to Stuyvesant because of the math and science," Kelley once pronounced, "but then we discovered that the worst teaching was in those subjects."

Kelley often needlessly antagonized teachers and administrators. Nevertheless, I credit him with going directly to the heart of the matter on the subject of the union contract. At one meeting the student delegates pushed for teachers to set aside time during the day to meet with students—in the same way that college professors are required to have office hours. This was an eminently reasonable request, since Stuyvesant teachers were notorious for leaving the building as soon as their last teaching period was over. Many even arranged their schedules so that they worked a five-hour day or less.

Suecoff responded that the proposal was a non-starter because all the hours of the school day allowed by the contract were already accounted for. A visibly frustrated Kelley then blurted out, in the truest words I had ever heard in the SLT meetings: "Every time the students try to bring up a program to improve education, we are told it violates the teachers' contract. It seems like education improvement is itself a violation of the contract."

–Sol Stern is a contributing editor at City Journal. *This essay is adapted from* Breaking Free: Public School Lessons and the Imperative of School Choice *(Encounter Books, 2003).*


NYPD
Courtesy
Professionalism
Respect
SCHOOL SAFETY
CS 6246

Jamaica High School in New York.

SECURITY DETAIL

An inside look at safety and discipline in the hyperlegalized world of a New York City high school

by MARC A. EPSTEIN

PHOTOGRAPHS BY PATRICK HARBRON

AT THE HEIGHT OF THE BABY boom, Jamaica High School in Queens, New York, enrolled approximately 5,000 students, who attended school in triple session. Each cohort of students had to start and end school at a different time in order to maximize classroom use. There were only two or three deans to address disciplinary issues, and a lone patrolman stood watch in the lobby of the school. It was assumed that a large portion of the school's graduates would go on to college, while the rest would become society's secretaries, plumbers, carpenters, printers, and butchers. Although Jamaica High was considered one of the school system's academic gems, it also provided vocational training that has long since disappeared.

Today the school population has been sliced to 2,500, yet Jamaica High now has eight deans (myself included), who devote much of their time to disciplinary issues; an assistant principal for security; two secretaries, one part time, one full time; and a school aide assigned just to the dean's office. In addition, ten school security agents employed by the New York City Police Department patrol Jamaica's halls. The number increases on those days when "random scanning" of students for weapons is in effect.

Schools Under Siege

The extra personnel alone do not even begin to account for the costs of attempting to maintain discipline and safety in the school. Consider the following tales from Jamaica High. In one instance, a teacher saw a student take a gun from his locker and show it to other students. He informed the security agent, who in dereliction of his duty did nothing. Then the teacher came to the Dean's Office, and I immediately began looking for the student. But the student had fled the building. When he returned two hours later, the police searched him and found nothing. The superintendent was informed, and an investigation of the event proceeded. The student claimed that he had no gun. When asked why he had exited the building, he claimed that a security agent had thrown him out. He could not identify the agent by sex or by race.

Statements were taken from students who had seen the gun, the student's parents were notified, and a hearing date was set. For the teacher and me to attend the hearing, substitutes had to be hired to cover our classes. At the hearing, after waiting for two hours, we received word that the parent had asked for a postponement by making a telephone call, which was within his rights. It was the end of the school year, so the case was carried over to the next fall. The postponed hearing was scheduled for September 13, which wound up being just two days after the horrible events of September 11. As I walked to the offices located near Madison Avenue and 22nd Street in downtown Manhattan, an acrid stench hung in the air, even though I was some distance from Ground Zero.

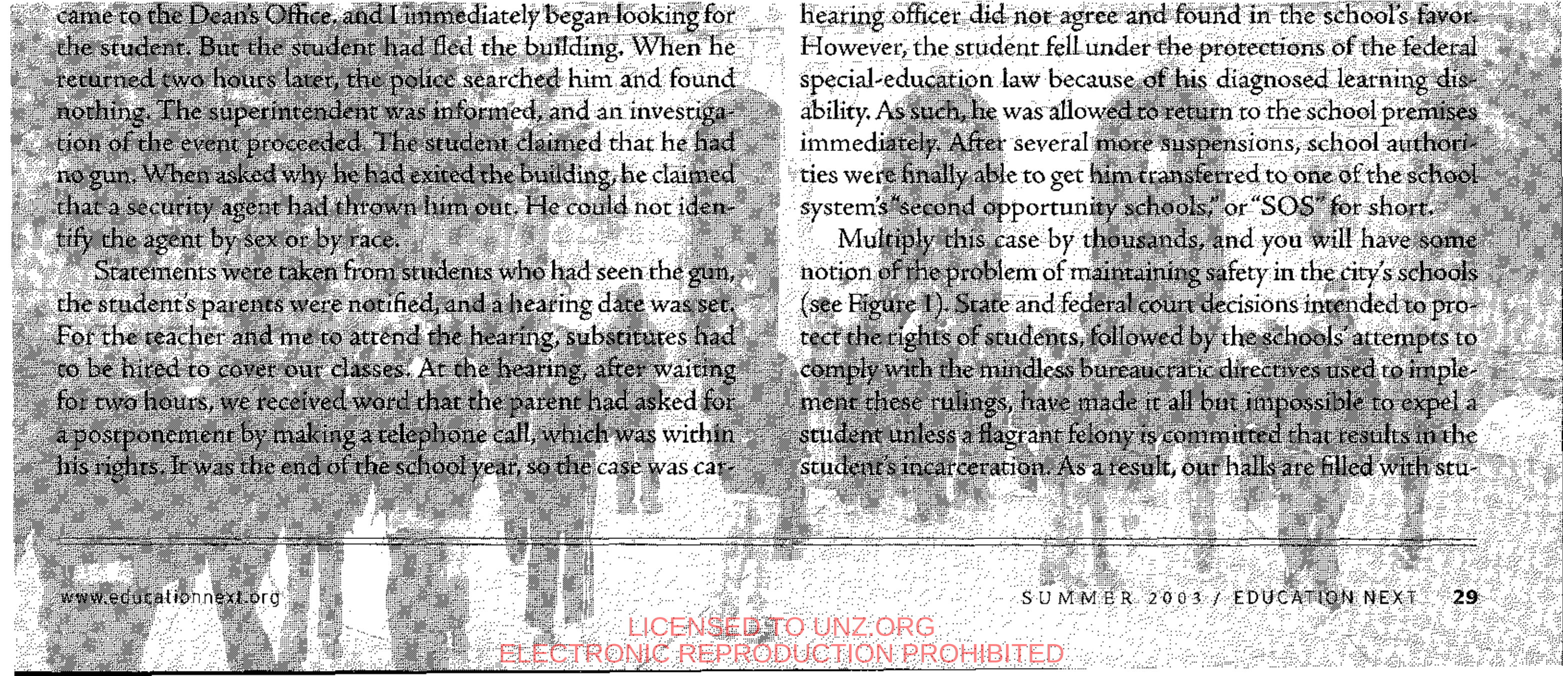
At the hearing, the parent of the accused asked me a very good question: Why had we allowed his son back to school if we believed that he had a gun? His son had rights, I said. And it took a couple thousand dollars' worth of personnel hours to protect them. The hearing officer eventually found in our favor, and the student was transferred to another school.

In another instance, a learning-disabled student in a mainstream class asked a teacher for a hall pass. When the teacher refused to give it to him, the student called the teacher, who had badly injured her hand in a car accident, a "crippled bitch" and threw an object at her. A security agent who had witnessed the event removed the student. The superintendent was informed, and the student's home was contacted. The next day I confiscated the hand of a store mannequin that the student was carrying around with him. He said he wanted to show the teacher that "he had a hand like hers." At 16 years of age, the student had earned just one of the 40 credits needed to graduate.

The teacher, the security agent, and I had to appear at the hearing. Again, substitutes were provided to cover our absences. And again, the hearing was postponed, this time at the request of a man who called himself the student's "grandfather." Even though he turned out not to be the grandfather, the hearing officer accepted him as the advocate because he brought a letter to that effect signed by the mother, with whom I was never able to speak.

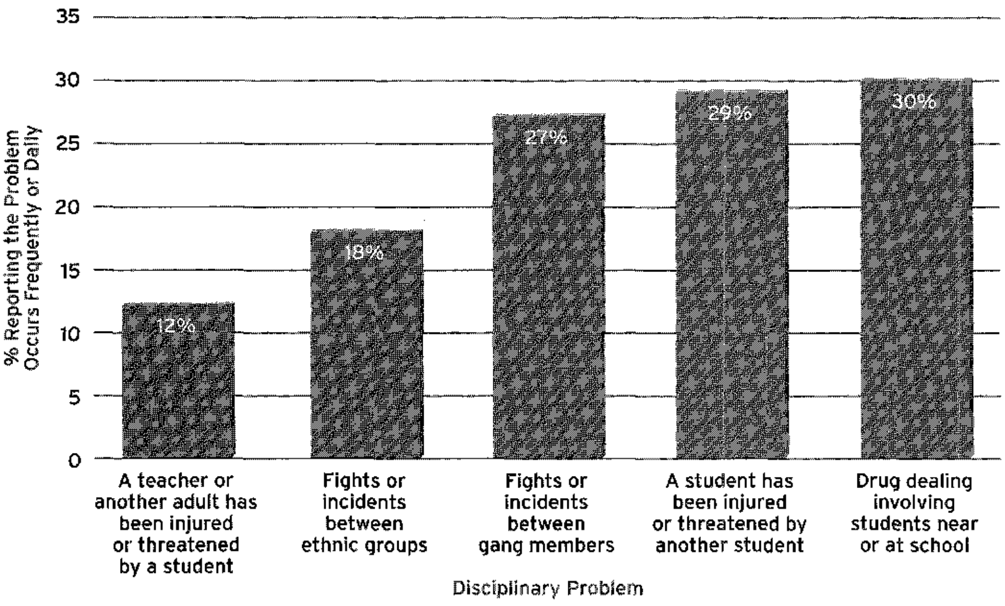
At the next hearing, the "grandfather" asked the teacher, "Do you consider yourself crippled, because if you do, then you are in fact a 'crippled bitch,' and the charges should be dropped." The hearing officer did not agree and found in the school's favor. However, the student fell under the protections of the federal special-education law because of his diagnosed learning disability. As such, he was allowed to return to the school premises immediately. After several more suspensions, school authorities were finally able to get him transferred to one of the school system's "second opportunity schools," or "SOS" for short.

Multiply this case by thousands, and you will have some notion of the problem of maintaining safety in the city's schools (see Figure 1). State and federal court decisions intended to protect the rights of students, followed by the schools' attempts to comply with the mindless bureaucratic directives used to implement these rulings, have made it all but impossible to expel a student unless a flagrant felony is committed that results in the student's incarceration. As a result, our halls are filled with stu-



Who's in Charge Here? (Figure 1)

A survey of New York City's public high-school students found that serious disciplinary problems occur in school frequently or daily.



SOURCE: Citizens' Committee for Children of New York, 2001

dents who cut nearly all their classes on a daily basis, whose only contribution to the educational process is to harass other students and make it necessary to maintain a phalanx of security officers at the school.

At a large high school like Jamaica, the demands of the school day make it nearly impossible to give these deeply troubled kids the kind of individual attention and support they need. They slip through the cracks and wind up disrupting not only their own education, but also that of the students around them. The schools need to be made safe for education, and those who bring chaos to the school environment need alternative education settings, where teachers and counselors can work with them and their families on a more one-to-one basis. This is especially true of students who have committed criminal acts in or outside of school. The New York City Department of Education took a step in the right direction in February 2003, when officials announced plans to open 17 "New Beginnings" centers for students with the most serious behavior problems. However, these centers will accommodate only a fraction of the students who require placement, and red tape has delayed their opening.

[At a large school like Jamaica High, everyday demands make it nearly impossible to give troubled kids the kind of support they need.]

Often these children can be targeted for help merely on the basis of their academic progress, but they rarely are. Our worst discipline problems usually present our biggest educational challenges as well. In almost every instance, an investigation of a serious disciplinary infraction reveals that the student has made little or no academic progress. For students who enter high school unable to read, the game is up unless they are identified and placed in an alternative learning environment outside of the high-school setting. Once inside the large New York City high school, they become lost to everyone except the police and the deans.

Fuzzy Math

It is impossible to account for the pedagogical costs such students exact on the schools: the classroom disruptions, the harassing of other students, and the frustrated new teachers who leave the profession. And the New York City Department of Education has not even attempted to tally the financial costs. Indeed, nobody knows how much of the city's $12 billion-plus annual budget is dedicated to school safety. But we do know that a significant portion of the funds that are earmarked for instruction and administration is actually used to maintain security.

There are approximately 150 high schools in the New York City school system, and a significant share of them house more than 2,000 students. High schools increasingly make security the sole responsibility of an assistant principal, which is supposed to free the school's principal to focus on instructional leadership. The responsibility for the day-to-day enforcement of each school's disciplinary code falls to the dean's office. School deans, whose number varies from school to school, usually teach only two classes a day. Yet at Jamaica High, the salaries and benefits for the eight deans and three-person support staff assigned to school security are treated as part of the instructional and administrative budget. Another 14 school aides are stationed in the cafeteria, outside bathrooms, locker rooms, and halls, in what can certainly be described as security-related assignments. At least $1 million in salary and benefits is devoted to school safety at Jamaica High alone—and that does not include the cost of the ten school safety officers, who are paid through a separate contract with the Police Department.

The situation is routinely aggravated by administrative decisions that fail to recognize the realities of everyday school life. For instance, schools can administer one of two kinds of suspensions, the "principal's suspension" and the more serious "superintendent's suspension." However, the principal's suspension was rendered essentially useless when the previous schools chancellor, Harold Levy, ordered that the suspension be served "in house." This was abetted by new state regulations requiring that suspended students not have their education interrupted.

POLICE
PATROL

A sizable portion of the student body regularly cuts class and keeps school safety and the academic deans at bay on a full-time basis.

PHOTOGRAPH FROM AP WORLD WIDE

Before Levy changed the rules, the principal's suspension was an effective tool because it kept an unruly student out of the building until at least one of the student's parents accompanied the student to a conference with the principal, the dean, and the guidance counselor. Students prone to misbehaving would think twice if their mother or father had to take off from work to get them readmitted. It also gave school administrators the opportunity to discuss the incident, and usually the student's general lack of academic progress, with the parent. The same conference is supposed to take place after an in-house suspension, but with the student already in school, the parent would often refuse to show up. The principal's suspension effectively withered as our first line of defense.

The granting of a superintendent's suspension hearing allows the student charged with an offense to call and cross-examine witnesses, postpone, and of course appeal the decision. The hearing is held in a special facility with a hearing officer, the dean of record, the victim or victims, witnesses, and a technician who transcribes the testimony. Before the New York state commissioner of education forbade it, students who had charges sustained against them in such a hearing would be moved to another school automatically. Under current policy, if the school believes the student represents a serious threat to the school population, it must initiate an involuntary transfer procedure, involving another round of reports and hearings before the student can be removed. This is essentially a game of musical chairs, only with a guarantee that when the music stops everyone is left with a chair. Should the case involve a "special needs" student, a finding must be made that the offense was not related to the disability. If the disability is ruled to have been a contributing factor, the student is returned to school without consequence.

For instance, one student was hearing voices and stalked a teacher, who had an order of protection issued. The student turned out to be schizophrenic and was removed as an EDP (emotionally disturbed person) on four separate occasions over two years. The father refused to have him institutionalized, despite the recommendation of the school psychologist. He was readmitted again and again. He eventually left the school, much as hundreds of students do each year, just disappearing into the ether without our ever knowing where they wound up. The point is that sizable populations in need of close supervision for a host of problems that often have nothing to do with education are simply left adrift in large city high schools. When their behavior becomes truly disruptive, the system responds with a stream of procedures that expends endless man-hours and appropriates uncountable treasure, often without anything to show for the effort.

The Hard Cases

In dealing with the more dangerous members of our student population, the system attempts to ensure safety by disrupting their pattern of behavior. Expulsion is no longer an option, so this is accomplished by using the lengthy involuntary transfer procedure to move the disruptive students to other locations. Eventually, however, cliques of students involuntarily transferred to other schools meet up once again. One is reminded of the film *Escape from Alcatraz*, where Clint Eastwood reconnects with his old cellmates from Joliet and plans his escape from "the Rock" along with them. Recently, the mother of a student who was new to the school called to tell me that her child had been robbed of her new leather coat after getting into an after-school fight with a group of students on Jamaica Avenue. The next day I asked the student, "How were you able to find yourself fighting with another group of students on your very first day?" She said that she had linked up with a bunch of fellow involuntary transfers from Franklin K. Lane High School, who began feuding with girls from Jamaica High. Why we would expect such students to perform any better simply by transferring them to yet another large high school is beyond me.

Since students must be given a fresh start, schools receiving transfers are not informed of the reasons why a new student has been placed in their midst. Even worse, the schools are not informed of arrests outside the school, so felons are admitted without our knowledge on a regular basis. Last year, I spent two months trying to track down a student who had committed a school infraction, only to learn that he had been discharged after 20 days of continuous absence. He was on Riker's Island, awaiting sentencing. Later, after being placed on probation for armed robbery, he was readmitted to Jamaica High. Court decisions and school regulations have ensured a safe haven for some truly dangerous criminals in our schools. Should a student be arrested in school, he can come back the very next day and await the results of school disciplinary action for several days afterward.

[Court decisions and school regulations have ensured a safe haven for some truly dangerous criminals in our schools.]

In a similar vein, the schools are mandated to educate children who have been removed from their homes by the courts and placed in group homes within our zone. Often they are the victims of the most horrific domestic circumstances. Placed in a school of 2,500 students, in the middle of a school term, what are the odds that they will successfully adjust? In November 2002, the director of a group home informed me that three of her charges had robbed another student and asked me if I wanted them arrested. When I asked her why they were placed at Jamaica, she informed me that there were no other options.

At any given time a high school of 2,000 or more students will have 300 to 400 students who have made no academic progress over the course of two years. A percentage will be passive and simply go to class and fail most of their courses. But a sizable portion will cut class regularly and keep school safety and the deans at bay on a full-time basis. With no high-quality alternative settings to send them to and no power to eject them from the premises, the school safety officers find themselves either herding them into classrooms or bonding with the worst offenders in a sort of reverse "Stockholm syndrome."

Snowballing

The loss of building control did not happen overnight. Decisions made on a case-by-case basis usually made sense, but when compounded with other "plausible" solutions, they incubated like one of those mysterious building molds that render a structure uninhabitable. For example, high schools used to have a homeroom at the start and end of the school day. On entry, students would lock up their coats and be marked present. At the end of the day, they would report to the room again and get their coats. Especially during the winter, this would help to control the building. Students could not simply walk out if they felt like it or flee if they were in trouble. School authorities could readily spot an intruder wearing a coat while everyone else was in shirtsleeves. It was also a good way to eliminate the problems of coats' being stolen and endless confrontations over the wearing of hats and gang-related paraphernalia. During the 1990s, however, the city's teacher union, the United Federation of Teachers (UFT), argued that forcing teachers to supervise homeroom was unprofessional. The union was able to get this effective building-control tool abolished through contract negotiations.

Take another example. School aides were put in charge of supervising the cafeteria after the UFT successfully argued that teachers should not have cafeteria duty. Even some of the severest critics of the UFT agreed that it was a waste of teach-

At least $1 million is devoted to school safety at Jamaica High—and that does not include the cost of the ten school safety officers, who are paid through the Police Department.

ers' time for them to be policing the cafeteria. However, teachers knew the students, and the students *knew that they* knew them. The teachers could keep students in line in a way that school aides never could.

In one fell swoop, the cafeteria became student-gang turf. In Park West High School, a separate cafeteria had to be created for freshmen because they were falling prey to upperclassmen. Naturally, the cafeteria has become a fertile breeding ground for conflict and disorderly behavior that bubbles up daily throughout the school. Students had at least a dose of respect for the school deans and teachers who used to manage the cafeteria. School aides, by contrast, are easily intimidated by students. It is a rare school aide who will confront gang-affiliated students without fear of retribution.

Not only the teachers but also frequently the parents of disruptive students are shocked by the laxity of the school system. My first superintendent's suspension hearing involved a freshman accused of threatening a female teacher who simply asked him to take his seat and stop cursing. He responded by overturning a desk as he walked toward her, stating he was "going to f--- her up." He was charged with threatening and menacing. The charges against him were sustained, and afterward his mother decided to return him to Trinidad, where the young man was from. When she returned to school to retrieve his records, she informed me that in Trinidad, her son would never have dared to threaten a teacher, and if he did . . . She did not complete the sentence. I would experience this sentiment time and again. Parental disbelief at the level of tolerance for this kind of behavior is palpable. Yet the only response from policymakers and the courts has been the erection of a costly system of adjudicating, tolerating, excusing, and ultimately ignoring deviant behavior.

Those strolling past the library at Jamaica High School can view the plaques honoring the school's alumni who gave their lives for their country during World War II. Pictures of students, faculty, and athletic teams also adorn the walls. The picture-taking stops sometime during the mid-1970s though. It is as if a geological age had been preserved in a stratified layer of sediment. There was a time, it seems, when the school had a culture and a sense of spirit that invited displays of unity and pride. Something quite meaningful appears to have been lost in the decades since.

Statistically, Jamaica High School falls squarely in the middle of the school system. Learning takes place, and good kids get educated by good teachers, but the presence of a permanent class of "students" who cannot function in this setting continually destabilizes the learning environment, tempts marginal students to behave badly, and squanders untold treasure for no good reason.

–Marc A. Epstein teaches history and serves as an academic dean at Jamaica High School in Queens, New York.

Philosopher or KING?

The ideas and strategy of legendary AFT leader Albert Shanker

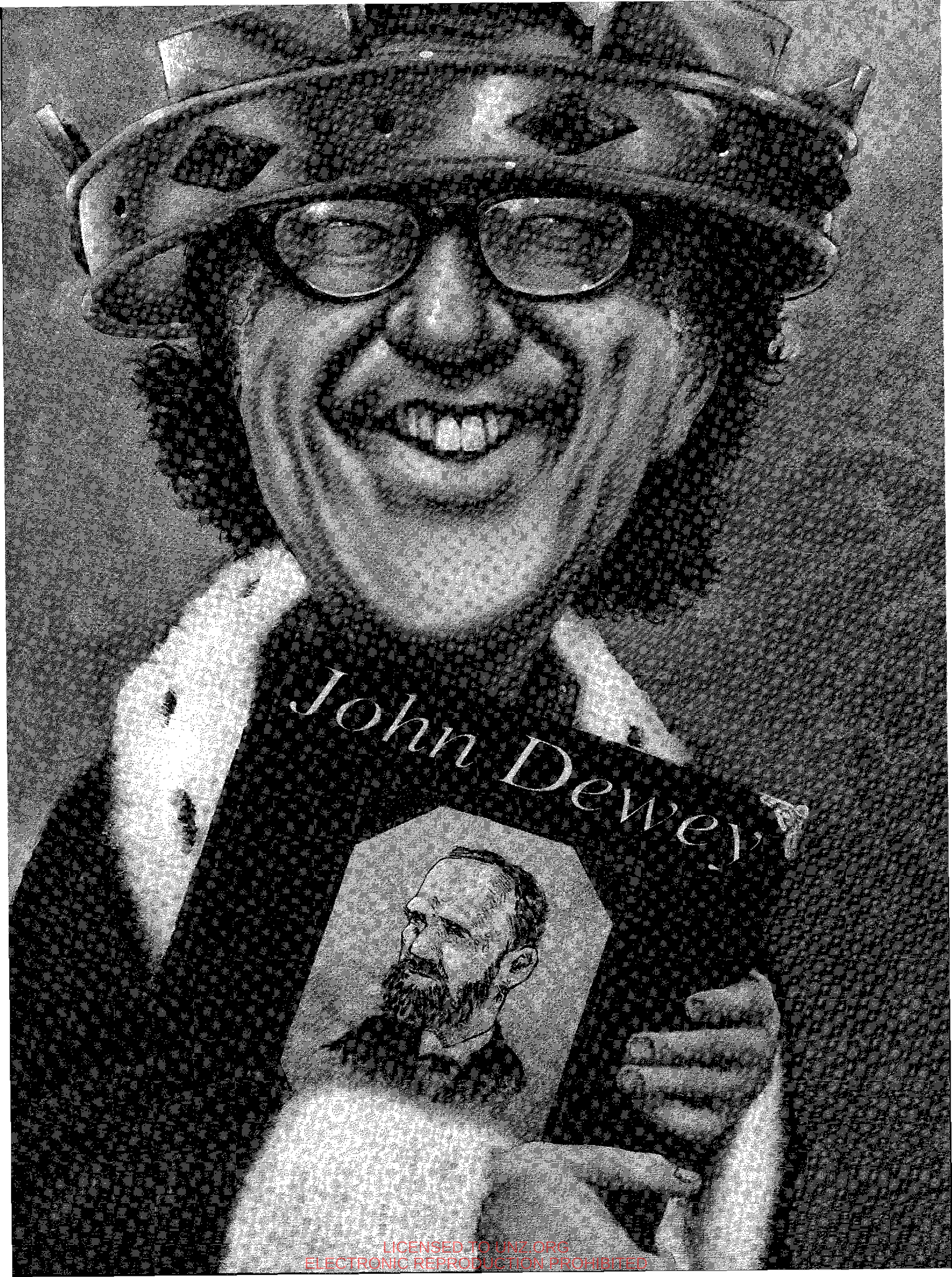
by RICHARD D. KAHLENBERG

THE MOST INFLUENTIAL VOICES IN AMERICAN EDUCATION during the 19th century and the first half of the 20th century were surely Horace Mann and John Dewey, respectively. But who ranks as the most consequential figure of the past 50 years? One of the many leaders who claimed the mantle of "education president"? A particularly influential U.S. secretary of education? A renowned education theorist? A pathbreaking state education official or schools superintendent?

Consider, instead, the late Albert Shanker—the president of a teacher union, of all things, and not even the largest one at that. Writing in the *New Republic*, Sara Mosle called Shanker, the legendary head of the American Federation of Teachers (AFT) from 1974 to 1997, "our Dewey, the most important American educator in half a century." The notion is astonishing—akin to claiming that the president of the United Auto Workers was responsible for the most important developments in the American automobile industry—yet well justified.

As one of the founding fathers of teacher unionism in New York City during the early 1960s, Shanker helped to create a movement that has become an enormous, if not the dominant, force in K–12 public education. During the 1980s and 1990s, he sought to transform teacher unions into a powerful voice for education reform, proposing ideas that were highly unconventional for a union president. In fact, the modern

ILLUSTRATION BY ZACHARY PULLEN


John Dewey

accountability movement, right through to the federal No Child Left Behind Act of 2001, owes much to Shanker's relentless calls for higher standards, assessments, and consequences for poor performance. Shanker was also an early proponent of public school choice, charter schools (some even credit him with the idea), rigorous knowledge and skills testing for teachers, and extra pay for master teachers. His support for these reforms sharply distinguished him from the leadership of the nation's largest teacher union, the National Education Association (NEA). By openly acknowledging the shortcomings of public schools and embracing innovation, he became a much more credible and effective voice for public education than the NEA or other defenders of the status quo.

Bipartisan Praise

Shanker's effectiveness as a leader stemmed from his unique combination of gifts. He was a union leader who could quote Aristotle and an intellectual who knew how to run a meeting. His innovative thinking drew invitations to teach from Hunter College and Harvard. Shanker was asked to join the boards of the Spencer Foundation and the Twentieth Century Fund, "one of only a handful of labor leaders to serve as a foundation trustee," one expert noted. At the same time, Shanker transformed the AFT, *Education Week* marveled, "into a labor union that often acts like a think tank."

He did not always come up with original ideas, but he took good ones and spread them through the 1,300 "Where We Stand" columns that he wrote in his lifetime; the column, a paid advertisement now written by Shanker's successor, still runs in the "Week in Review" section of the Sunday *New York Times*. "The impact was extraordinary," said the late U.S. senator Daniel Patrick Moynihan. "Union leaders in those days rarely wrote essays, still less felicitous, thoughtful analysis of public policies." Today every leader with some institutional money, including the presidents of the NEA and the AFT, seems to have a paid column, but none reads like Shanker's "Where We Stand." Shanker's intellect and skills as a debater were backed by a forceful personality and real-world power. His power, in turn, was bolstered by longevity: while he ruled the AFT for 23 straight years, the president of the NEA is term-limited.

Shanker sought to transform teacher unions into a powerful voice for education reform, proposing ideas that were unconventional for a union president.

Shanker gained a hearing for his ideas from both liberals and conservatives, not by taking moderate positions that consistently split the difference, but by embracing a coherent philosophy that sometimes led to "liberal" policy conclusions, other times to "conservative" ones. Shanker allied with liberals on trade unionism, public schools, and economic equality, while finding common cause with conservatives on issues like standards, public school choice, racial preferences, bilingual education, and communism. At the height of the war between Ronald Reagan and organized labor, the AFT invited the president to speak at its convention, and Reagan accepted.

To be sure, Shanker's legacy is still a matter of some debate. Critics contend that Shanker's words and his union's actions were often worlds apart, particularly during Shanker's later years as president of the AFT. For instance, in his 1997 book *The Teacher Unions*, Myron Lieberman argued, "The neoconservative notion that the AFT is a more enlightened union or more hospitable to educational reform or innovation [than the NEA] resulted from AFT president Albert Shanker's ability to manipulate media, not to any substantive differences between the unions."

While it is true that the AFT's local unions, locked in hardball urban political struggles and pressing mainly bread-and-butter issues, often strayed from Shanker's bold vision, his leadership on a variety of issues exerted broad influence—mostly for the better—on the course of education reform. His weekly "Where We Stand" columns did so much more than merely cheerlead for public education and nurture the mom-and-apple-pie view of teachers, which is what one might expect from an advertorial run by a labor union. Shanker was as apt to criticize public schools as he was to condemn those who would eviscerate them, as likely to call for higher standards for teachers as he was to protect their tenure rights. His influence can still be felt most palpably in four areas: teacher unionism; the standards movement; the choice movement; and the adoption of reforms to improve teacher quality.

Organization Man

Until the fledgling United Federation of Teachers (UFT) in

New York City won collective-bargaining rights in 1961, teacher unions were impotent. It had been considered pointless to organize teachers, or other public employees for that matter, because it generally was (and is) illegal for public servants to go on strike. The NEA, then a loose professional association, opposed collective bargaining as "unprofessional" and fought hard against the principle in New York City and elsewhere.

Shanker, an official with the UFT during the 1950s and 1960s, and his colleagues at the AFT sought to test those assumptions. Shanker had been a Ph.D. candidate in philosophy at Columbia University, but he quit for lack of funds and became a substitute teacher in the New York City public schools, where he had once been a star debater at the elite Stuyvesant High School. He soon became a union organizer and, in the years to follow, he would serve two jail sentences for his unlawful strike activities. Militancy worked, however, and teacher unionization spread rapidly: the AFT signed up thousands of new members. This forced the much larger NEA to change its position on collective bargaining. Now teachers are the most unionized sector in the workforce. While trade unionism was in general decline, the AFT grew from 71,000 members in 1961 to close to one million at the time of Shanker's death in 1997. He was, wrote Mosle, "one of [the 20th] century's greatest labor organizers." The UFT's example helped to spark unionization among other civil servants, from about 5 percent of public workers in the early 1960s to 43 percent today. At Shanker's memorial service in 1997, Daniel Patrick Moynihan remarked, "If, as Emerson wrote, an institution is the lengthened shadow of one man, the public employees unions are singularly the mark of Al Shanker's inspired life."

Shanker acknowledged that the goals of teacher unions and the needs of students and effective schools could sometimes come into conflict. Initially he focused on traditional union activities: building the membership and securing better salaries, benefits, and protections for teachers. But as the AFT and NEA grew entrenched and powerful, Shanker attempted to recast teacher unionism as a force that looked beyond teachers' narrow interests. Along the way, he transformed his image from union hothead to educational statesman. Shanker's key insight was that he could motivate his members to push for better public schools by holding up the threat of school vouchers.

Shanker realized that the transformation of vouchers from a theoretical idea proposed by conservative economists in the 1950s to a viable political threat by the late 1970s gave union leaders the opportunity to engage in public school reform as never before. It was not only an opportunity, but a necessity. After all, if the public schools continued to underperform, the nation's confidence in public education itself—which to Shanker was both a positive democratic force in society and the root of union leverage—would dissipate. Shanker thundered at union delegates: "It is as much your duty to preserve public education as it is to negotiate a good contract."

Setting the Standard

As part of his reformist message, Shanker became a leading promoter of the standards movement. Indeed, no one is more responsible for its emergence than Albert Shanker. Today standards lie at the core of education reform packages put forth by presidents as ideologically disparate as George W. Bush and Bill Clinton, but the idea was not always so widely accepted. For years, standards were opposed by a coalition of liberals and conservatives, though for different reasons. Liberals worried that poor and minority students would be penalized by high-stakes tests, while conservatives wanted to preserve the "local control" of schools.

The turning point came in 1983, when the Reagan administration's National Commission on Excellence in Education issued the *A Nation at Risk* report (see the Spring 2003 issue of *Education Next* for a retrospective on the *Risk* report). The report bemoaned the nation's declining test scores and the absence of a serious curriculum in American public schools. While the NEA and the rest of the education establishment lined up to assail the report, AFT president Shanker shocked observers by endorsing the report's findings. In so doing, he helped to secure *Risk*'s lasting influence. "It was vital that someone with stature step up," said Milton Goldberg, the commission's executive director. "Al Shanker never wavered on that issue, and the rest of the education community and public finally caught up to him."

> While the NEA and the rest of the education establishment lined up to assail *A Nation at Risk,* AFT president Shanker shocked observers by endorsing the report.

Without Shanker, the AFT was likely to have joined the opposition, potentially crippling the report. Current AFT president Sandra Feldman explained to journalist Sara Mosle that when advance copies of *Risk* were circulated on the eve of a union convention in 1983, "We all had this visceral reaction to it. You know, 'This is horrible. They're attacking teachers.' Everyone was watching Al to hear his response. When Al finished reading the report, he closed the book and looked up at all of us and said, 'The report is right, and not only that, we should say that before our members.' And that's what he did. It was a really courageous thing for him to do." The AFT had always backed tough disciplinary standards (which was clearly in the self-interest of teachers), but the new focus on academics would require much more of, and place greater pressure on, rank-and-file teachers.

A Nation at Risk helped to launch the modern education reform movement. Beginning in the late 1980s, Shanker pushed hard for state and federal legislation to raise academic standards, and he kept the pressure on for educational testing and consequences for poor performance. He reminded officials that when he was a teacher and gave an exam or quiz, the students always asked, "Does it count?" During the 1994 reauthorization of the federal Elementary and Secondary Education Act, which ultimately forced the states to begin developing their accountability systems, Shanker was a staunch proponent of tough standards, and penned a pivotal article blasting a proposal to water down the bill. After President Clinton devoted much of his 1997 State of the Union address to standards, he called the then-ailing Shanker and told him, "Al, this is your agenda. We should have listened to you sooner."

Whereas the NEA leadership supported bilingual education and racial preferences, Shanker raised objections to both, in part because he thought they held minority students to lower standards. Even today, the AFT continues to lend some support to standards-based reform, while the NEA has opposed high-stakes testing, supported parents who boycott tests, and funded researchers to try to undermine standards-based reform.

Shanker was committed to a strong union movement for the same reason he was a strong cold warrior and a promoter of common standards: in his view, they all served democratic values.

Freedom with Accountability

Shanker later made an inspired link between his push for standards and the nascent idea of charter schools, publicly funded schools that would operate independently of school districts but be subject to tough performance requirements. In 1996, Shanker told Mosle, "If we had a system of standards and assessments in place, then as far as I'm concerned every school should be a charter school." In Shanker's view, these schools should be set up by groups of teachers and parents who wanted to try innovative educational approaches. This vision differs somewhat from that of the management of charter schools by privately run corporations, a trend that the AFT has opposed in recent years.

Shanker began championing charter schools as early as 1988, though his contribution is a matter of some debate. Brookings Institution scholar Diane Ravitch says Shanker "initiated the idea of charter schools," while the authors of *Charter Schools in Action* credit Shanker with coining the term "charter schools" after visiting a school in Cologne, Germany. Others note that Shanker himself attributed the idea to a retired educator named Ray Budde and argue that Shanker's role was mainly to disseminate and popularize it.

Before charter schools, there were district-run magnet schools and other forms of public school choice, of which Shanker was an early proponent. In April 1985, he argued that "the greatest possible choice among public schools" should be provided as a way to promote a diversity of school offerings and to better match the interests of individual teachers and students. The key difference between public school choice and vouchers, he insisted, is that under public school choice, "everyone competes under the same ground rules."

With the NEA opposed to schools of choice and charter schools at the time, both reforms might have died without Shanker's support. Instead, they have become significant parts of the education landscape. Public school choice has exploded to include more than five million students, a number that will surely rise under the requirements of the federal No Child Left Behind Act. Charter schools served no students in 1990, but today they educate more than 700,000.

Alienating the Troops

Even more controversial among teachers than Shanker's advocacy of high standards and public school choice was his embrace of a series of reforms intended to improve the quality of the teaching profession.

To attract good teachers, conservatives had long argued for "merit pay" as a means of rewarding superior teachers. To maintain solidarity, however, teacher unions typically oppose paying different salaries to teachers who possess the same amounts of experience and education. Shanker offered a compromise. Instead of basing pay on the judgments of school administrators, which might be open to favoritism, Shanker called for a system of board certification, like that used for doctors. Shanker's idea, which was first outlined by teacher union gadfly Myron Lieberman, became the private, nonprofit National Board for Professional Teaching Standards. Many states and school districts now provide bonuses to teachers who obtain board certification. Former North Carolina governor Jim Hunt, a prime mover behind the idea of board certification, argues, "More than any other single person, Al Shanker was the founder of the [group]."

Shanker supported tenure protections for teachers after a three- or four-year probationary period because he believed that it was an important tool for attracting high-quality teachers and protecting academic freedom. But he argued that tenure should not protect bad teachers and that unions should be at the forefront of "removing those who are incompetent, with due process." In particular, Shanker supported a highly controversial experiment of the local union in Toledo, Ohio. The union had worked with the school district to institute peer review, a program where top-flight teachers evaluated new teachers and even veteran teachers who had severe problems. In 1982, Shanker invited the president of Toledo's teacher union, Dal Lawrence, to speak about his program to the AFT's executive council in Washington. Lawrence said he wasn't sure if he would be praised or reprimanded. "Here we were, a teachers union, and we were evaluating and even recommending the non-renewal and terminations of teachers," Lawrence recalls. After he gave his presentation, many council members were harshly critical. Then Shanker, who had remained silent for much of the discussion, told the council, "I think you're missing something here." He argued that other professions took responsibility for maintaining high standards among their members, and he endorsed the Toledo model.

Shanker argued that tenure should not protect bad teachers and that unions should be at the forefront of "removing those who are incompetent, with due process."

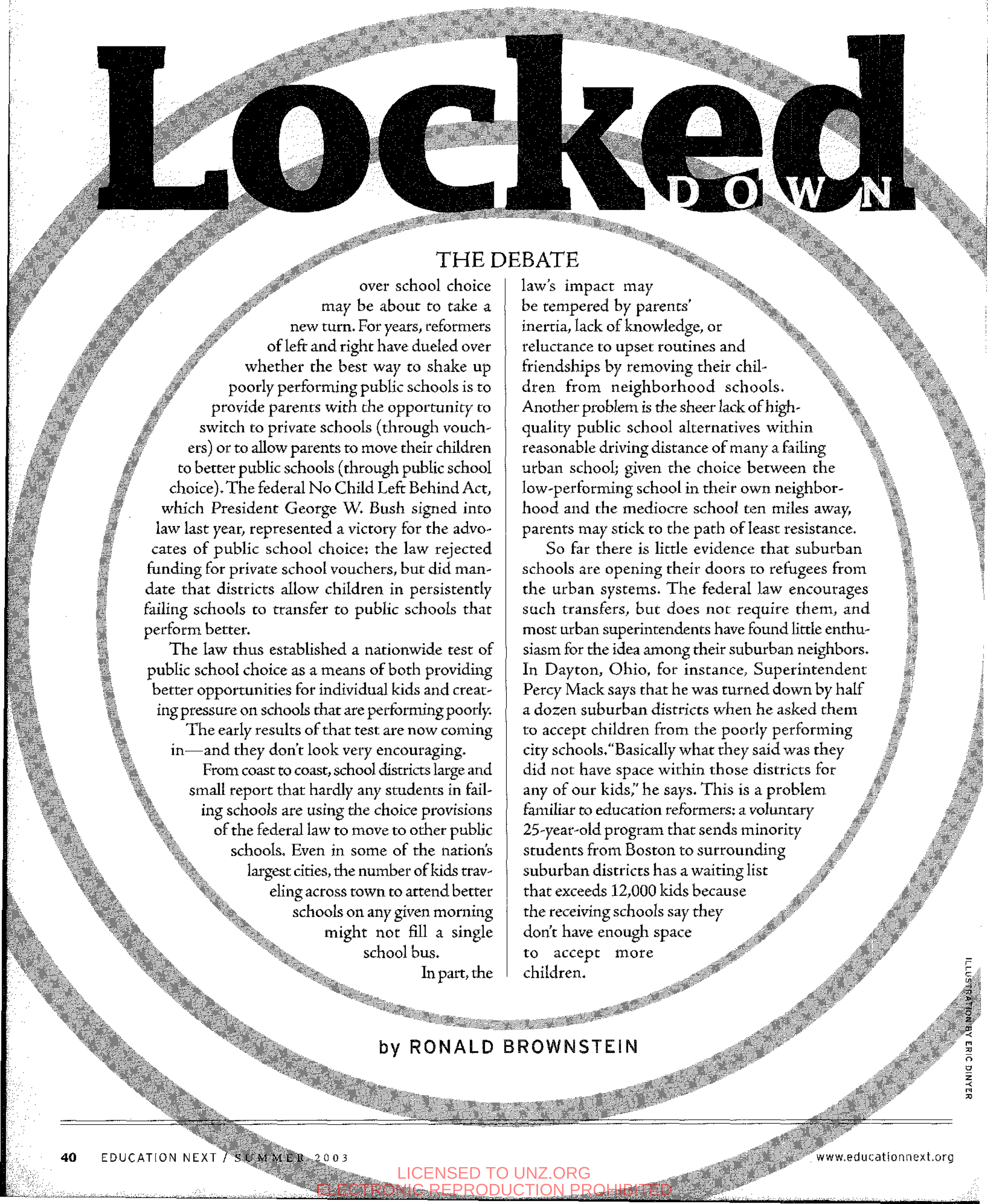
Likewise, Shanker shocked observers in 1985 by backing a rigorous National Teacher Competency Exam for new teachers, similar to that used by the legal and medical professions. The NEA opposed a national standard, arguing that state-by-state minimum requirements were sufficient. But Shanker derided existing state-level teacher tests as the equivalent of asking doctors to pass a test in "elementary biology." He backed up the proposal with a declaration that the AFT would limit membership to those who passed.

Progress on Shanker's trio of teacher-quality initiatives has been slow. In 1996, Shanker told journalist Thomas Toch, "Convincing people to change has been a damn difficult thing to do. I would go into a state, talk up reform, and as soon as I left, the union attorney would come in and say, 'We've got a great tenure law, let's keep it.'" Peer review is limited to about 50 sites across the nation, and many states still certify teachers with weak exams—or none at all. National board certification stands at 24,000 teachers. On these reforms, Shanker was more successful at stoking the dialogue than winning over the membership.

On the whole, however, Shanker's strategic vision paid off. By acknowledging the trouble with public schools and advocating radical reforms, he gave a positive voice to teachers' concerns, keeping them in touch with the larger political dialogue. The tactic of the education establishment, to routinely call "for more money to address problems that they frequently argued didn't exist," notes Toch, was far less credible than Shanker's approach: fighting "to change public education in order to preserve it." When conservatives said public schools lacked the high standards of private schools, Shanker agreed, and he pushed for reform. When they said it was unfair to trap poor kids in bad schools, he agreed, and he championed public school choice as a better alternative. When conservatives criticized bilingual education and extreme multiculturalism for fracturing the nation's social cohesion, the union leader concurred, fought those programs, and argued that private school vouchers would promote even further balkanization.

Shanker was committed to a strong union movement for the same reason he was a strong cold warrior and a promoter of common standards: in his view, they all served democratic values. In this sense, then, Albert Shanker had more in common with Horace Mann and John Dewey than just exerting a profound influence on American education.

–Richard D. Kahlenberg, a senior fellow at the Century Foundation, is writing a biography of Albert Shanker. His most recent book is All Together Now: Creating Middle-Class Schools through Public School Choice *(Brookings, 2001).*



Locked DOWN

THE DEBATE over school choice may be about to take a new turn. For years, reformers of left and right have dueled over whether the best way to shake up poorly performing public schools is to provide parents with the opportunity to switch to private schools (through vouchers) or to allow parents to move their children to better public schools (through public school choice). The federal No Child Left Behind Act, which President George W. Bush signed into law last year, represented a victory for the advocates of public school choice: the law rejected funding for private school vouchers, but did mandate that districts allow children in persistently failing schools to transfer to public schools that perform better.

The law thus established a nationwide test of public school choice as a means of both providing better opportunities for individual kids and creating pressure on schools that are performing poorly. The early results of that test are now coming in—and they don't look very encouraging.

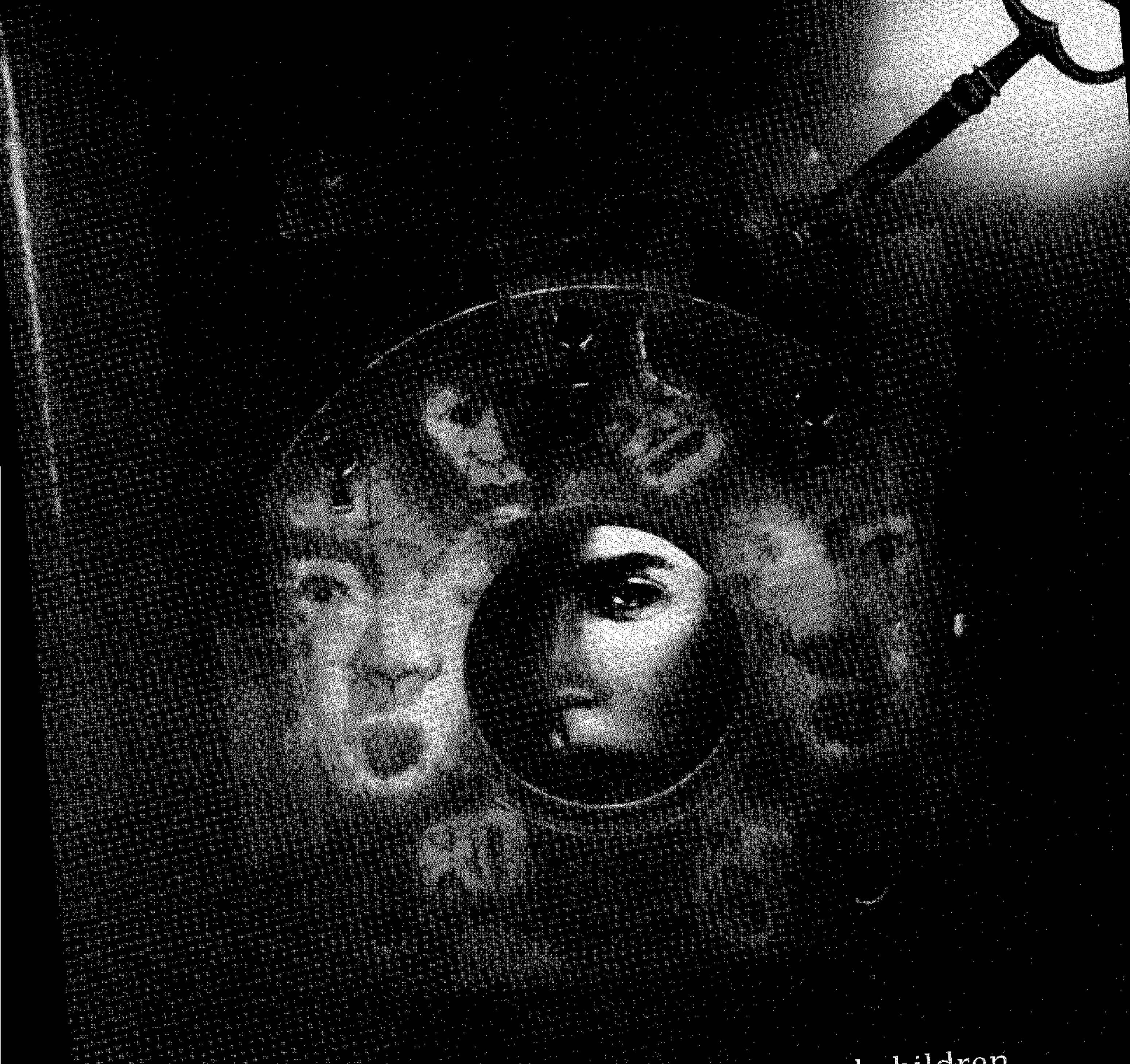
From coast to coast, school districts large and small report that hardly any students in failing schools are using the choice provisions of the federal law to move to other public schools. Even in some of the nation's largest cities, the number of kids traveling across town to attend better schools on any given morning might not fill a single school bus.

In part, the law's impact may be tempered by parents' inertia, lack of knowledge, or reluctance to upset routines and friendships by removing their children from neighborhood schools. Another problem is the sheer lack of high-quality public school alternatives within reasonable driving distance of many a failing urban school; given the choice between the low-performing school in their own neighborhood and the mediocre school ten miles away, parents may stick to the path of least resistance.

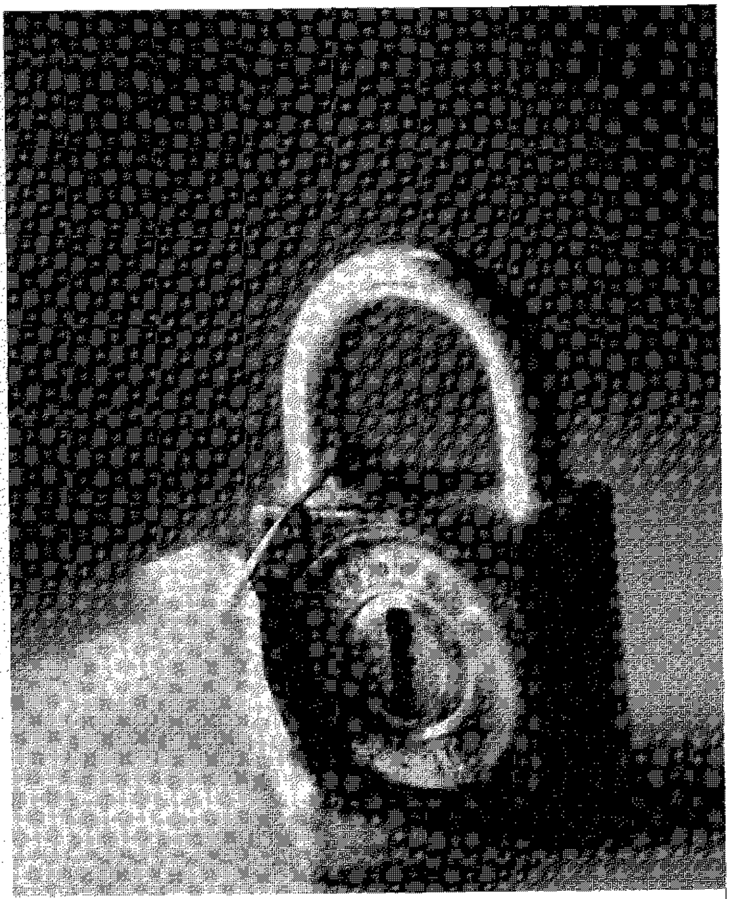
So far there is little evidence that suburban schools are opening their doors to refugees from the urban systems. The federal law encourages such transfers, but does not require them, and most urban superintendents have found little enthusiasm for the idea among their suburban neighbors. In Dayton, Ohio, for instance, Superintendent Percy Mack says that he was turned down by half a dozen suburban districts when he asked them to accept children from the poorly performing city schools. "Basically what they said was they did not have space within those districts for any of our kids," he says. This is a problem familiar to education reformers: a voluntary 25-year-old program that sends minority students from Boston to surrounding suburban districts has a waiting list that exceeds 12,000 kids because the receiving schools say they don't have enough space to accept more children.

by RONALD BROWNSTEIN

ILLUSTRATION BY ERIC DINYER



The No Child Left Behind Act granted children
the right to transfer out of failing schools.
let them?



These obstacles are compounded by the fact that few districts are making it easy for parents to exercise their right to choose or to avail themselves of the related option that offers "supplemental services," such as after-school tutoring, to students who remain in schools that have failed to improve student performance.

Massive resistance might be too strong a term to describe the way in which local school officials are implementing these new options for parents. But not by much. Using both subtle and overt strategies, school districts of every size have made it difficult for parents of children in failing schools even to learn about the new choices, and they have structured the programs in ways that make them less attractive to the parents who might be interested. "The only way you make something like this work is to fully inform parents what their options are and how to exercise their options, and school superintendents aren't doing that," says William L. Taylor, chairman of the Citizens Commission on Civil Rights, a liberal advocacy group.

The public school choice guarantees of the No Child Left Behind Act have had as much impact on the major school systems as a Ping-Pong ball fired at a battleship.

The lack of enthusiasm—and in some cases overt hostility—toward the new requirements underscores the difficulty of implementing any reform that requires school districts to impose changes that challenge their bureaucratic self-interest. It also raises questions about whether public school choice, as presently constructed, can have anywhere near the impact its supporters have long hoped. The coming debate will be over whether the solution is to create a more sweeping form of public school choice or to revive private school vouchers to create the alternative the public system has so far squelched.

Resistance Movements

The public school choice and supplemental services provisions of the No Child Left Behind Act were to be the most tangible lifelines for parents whose children attend low-performing schools. Schools that fail to make "adequate yearly progress" in improving student performance on standardized reading and math tests for two years in a row are subject to the act's sanctions. Their students must be allowed to transfer to a better school, with the school district paying the transportation costs. Alternatively, children who choose to remain in low-performing schools are eligible for after-school and weekend tutoring once their school fails to make adequate yearly progress for three years' running.

This approach was meant both to widen opportunities for students and to place competitive pressures on the schools. Schools that didn't improve would risk losing students and the accompanying state financial aid. They would also be forced to divert some of their Title I money into providing transportation and after-school tutoring programs. This was supposed to give schools and districts powerful new incentives to improve.

Or to resist the law.

In its first year, the transfer provisions of the new federal education law have had as much impact on the operations of the major school systems as a Ping-Pong ball fired at a battleship. In Chicago, of the 125,000 kids in 179 failing schools who were eligible to transfer to other public schools last September, fewer than 800 have switched. In Los Angeles, where about 200,000 students in 120 schools were eligible, fewer than 50 have changed schools. In New York, where 220,000 children in more than 300 schools were eligible, just 1,507 moved.

It's not only in the largest cities where the law has fizzled. In Cleveland, where 15,000 students in 21 schools were eligible, just 36 children requested transfers in the fall semester—and, of those, nine eventually returned to their original schools. In Boston, where students in 65 schools were eligible, apparently no students have used the new law's provisions to change schools. Likewise, no students have moved in Dayton, Ohio, though 10 of the district's 25 schools were on the state's list of failing schools. In Louisville, Kentucky, 2,900 kids in the Jefferson County Public Schools were eligible to transfer. Only 180 have moved.

What went wrong? The answer varies from city to city, though similar threads run through many of the tales.

In Chicago, local and state officials limited the program in ways that severely reduced its attractiveness to parents. Last summer, the state legislature passed a law saying that schools with selective enrollment—such as centers for gifted children—and those considered to be operating over capacity did not have to accept transfers from the poorly performing schools. (The federal law allows such exemptions as long as they were in place before July 1, 2002.)

Liberal education reformers argue that suburban school districts should be required to accept transfers from urban schools.

Then the district offered transfers only to students in 48 of the 179 schools that had failed to make adequate yearly progress. All of the 48 schools were elementary schools. In the end, even with these limits, about 2,000 parents requested transfers, says Phil Hansen, the chief accountability officer at the Chicago Public Schools. But because of capacity constraints at the schools designated to receive the transfers, just 1,100 applications were granted. By the time school was under way, many parents had second thoughts, and fewer than 800 children had moved.

The Neighborhood School

For the city, Hansen says, the moral of the story was that most parents don't want to move their children from their neighborhood school, no matter how miserable its scores on standardized tests. "The lesson we learned, which we kind of knew already," he says, "is that Chicago is a city of neighborhoods. Parents take pride in their neighborhood school; even if it is a low-performing school, parents feel closeness to that neighborhood school. What we heard from parents more often than not is, I don't want my children to go to another school, but what are you going to do to make my school better?"

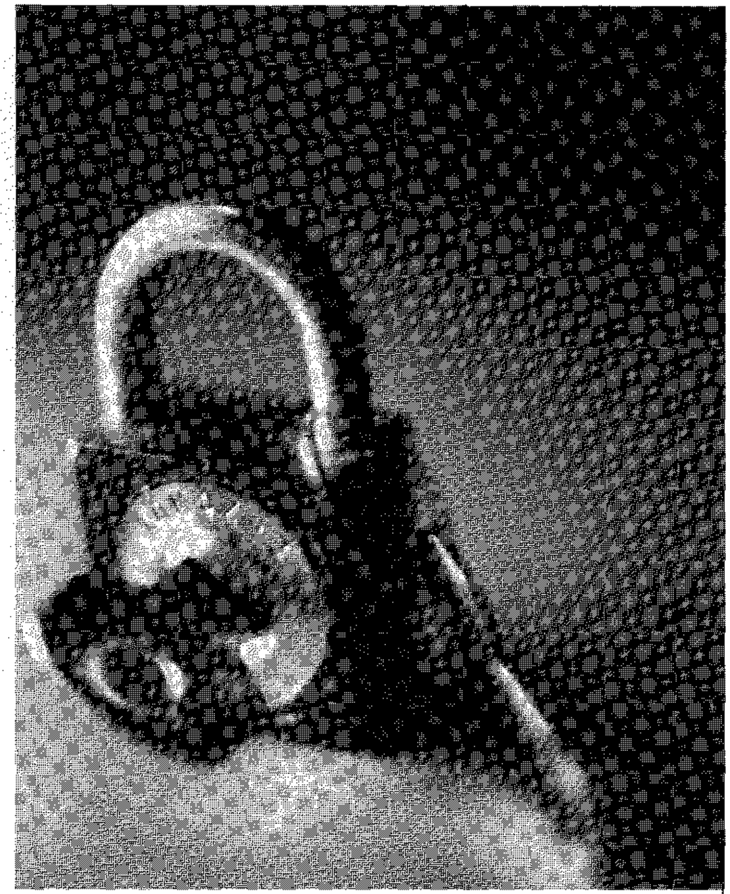
Madeline Talbott, the lead organizer for Illinois ACORN, a group that advocates on behalf of low-income families, doesn't entirely disagree. Parents, she says, are reluctant to send their children to a different neighborhood, one that they might have difficulty reaching in an emergency. But she says the way the city limited the choices gave parents little incentive to accept that risk.

"There were a lot of reasons we felt the way the thing was set up was pretty useless," she says. "We felt the choice provision, at least here, was something of a sham. It was a very expensive transfer of kids from one school to another that wouldn't necessarily be an improvement and would not necessarily have the support of the parent."

The problem began, she noted, when the state exempted many of the best schools in the city from accepting students from the low-performing schools. The way the federal law measures schools compounded the difficulty. Under the No Child Left Behind Act, schools are ranked by the trend, not the absolute level, of their students' performance on the standardized reading and math tests. That meant a school might be designated "needing improvement" because it had failed to raise scores over the past several years—but still could have a higher absolute score than a school that had met the federal standard because it had made steady gains from a lower base. The result was that Chicago, in some instances, offered parents a chance to transfer their children into schools whose overall test scores were lower than those of the schools they were already attending.

In the end, ACORN supported the city's proposal to limit choice to just the 48 schools and to use most of the money that would have gone into busing to fund improvements in the failing schools. "If we could have constructed a plan that gave children in low-income, low-performing schools a real opportunity, we would have supported it," Talbott says. "But we looked and looked and could not figure one out."

The lack of better alternatives for parents in poorly performing schools may be an even greater problem in smaller cities. In Dayton, Ohio, the initial state review concluded that all 25 of the district's schools had failed to make adequate progress. After further analysis, the state concluded that test scores had improved enough at 10 of the city's schools to lift them out of that category. "However, those 10 schools resembled the 15 that did need improvement," says schools superintendent Percy Mack. "Some of them even had lower scores"



than the 15 schools that remained on the state's list of failing schools.

In that circumstance, he says, the district decided that it made little sense to offer parents a chance to transfer to schools that were no better than the ones they were attending. It's a problem, he says, that is likely to be common in smaller and older cities. "In the major cities, where you are covering major areas of space, where some of the city schools extend out into the suburban areas, then you may find you have different looks and achievements in the schools," he says. "But in the smaller areas . . . you are going to see a lot of the schools look very similar."

The experience in Portland suggests that there may be a natural limit to how many parents will move their children from one conventional public school to another.

Didn't You Get the Notice?

In New York City, the deficiency was less in capacity than in motivation. "There is a lot of blame to go around," says Eva Moskowitz, a centrist Democrat from the east side of Manhattan who chairs the City Council's education committee. "I think on the local level there really is resistance to embracing choice."

The breakdown began, she says, with the New York City Department of Education's decision to let each school district in the city decide how to contact parents about the option to transfer from poorly performing schools. Rather than writing directly to parents, she says, many districts apparently sent the notice home in children's backpacks—an excellent way to ensure that few parents ever see it. And for those who managed to fish out the notice from the swamp of old homework assignments, baseball cards, and snack wrappers in the backpack, the letters themselves weren't much more illuminating.

"The letter was not a particularly encouraging letter, and it was quite difficult to understand," says Moskowitz. "I have a Ph.D. in American history and I had to read it about three times to figure out exactly whether this choice was guaranteed, and who do I contact, and am I going to have to pay for the transportation? It also wasn't clear if I could pick a school out of my district."

As a result, few parents seem to have known about their options. A December 2002 poll by the Foundation for Educational Reform and Accountability found that 85 percent of New York City parents with children in failing schools were unaware that the schools were on the state's list of low-performing schools. The Albany-based foundation, which supports school choice, also found that 94 percent of the parents were "likely" to request a transfer if they were made aware of the option.

Even parents who understood their rights then found themselves confronting the labyrinth of the city's often impenetrable school bureaucracy. As in Chicago, "Parents who wanted a choice complained that the choice they were given had worse scores than the school their child was in," Moskowitz continues. "Others found that . . . if they had questions, and they wanted to see the school [available for transfer] they were told they weren't allowed to visit. The stories run the gamut, but it does seem there was either passive or active resistance to offering parents real choice."

Similar problems later resurfaced in the way the school districts notified parents of students who were eligible for the new after-school tutoring services established under the federal law. After the *New York Times* revealed last

November that only 10,000 of the nearly quarter-million eligible students had signed up for the supplemental services, the city was forced to extend the deadline for applications. The problem, one city Education Department official acknowledged, is that districts did not want to lose the Title I money they would have to give to parents to obtain the after-school tutoring—and so did little to make parents aware of the opportunity. "Part of the problem is folks wanted to keep the money within the system," said the official. "Did the folks on the ground do an adequate job of saying to parents that the resources were available? The answer is no."

The breakdown in both the choice and the tutoring programs was so overwhelming that last December, New York City mayor Michael Bloomberg and Joel Klein, the new, Bloomberg-appointed schools chancellor, announced that they were taking over the process. Now, Bloomberg said, the central Department of Education would assume responsibility for notifying parents and establishing a process that would allow students in failing schools to transfer to better schools across the city, even outside their home districts. Watching the announcement, Moskowitz was encouraged but not entirely convinced. "I guess I am of two minds," she says. "It's hard to imagine it getting any worse.... Having said that, it's not as if the central [school administration] has any track record of doing this, or anything else, particularly well." New York City's intransigence even provoked a class-action lawsuit filed in January 2003 by parents claiming that the New York and Albany school districts had denied students their rights to transfer and to free tutoring.

Liberal education reformers argue that suburban school districts should be required to accept transfers from urban schools.

New York was hardly alone in cloaking the new options. One potential reason that so few students transferred in Cleveland, for instance, is that the district didn't notify parents that the choice was available until four days before school began—at which point understandably few were enthusiastic about uprooting their children. In Los Angeles some parents were not notified until after the school year began. The breakdowns extended up the bureaucratic line; many states didn't notify cities which schools were failing until late in the summer, which gave them little time to contact parents. And many states were late in designating the firms that could provide after-school tutoring, with the result that districts, once again, had less time to notify parents. Illinois education officials initially told the Chicago schools they would not have to offer the tutoring services in failing schools until September 2003—which left district officials scrambling when the federal Department of Education decreed that they would have to begin offering the services this past winter.

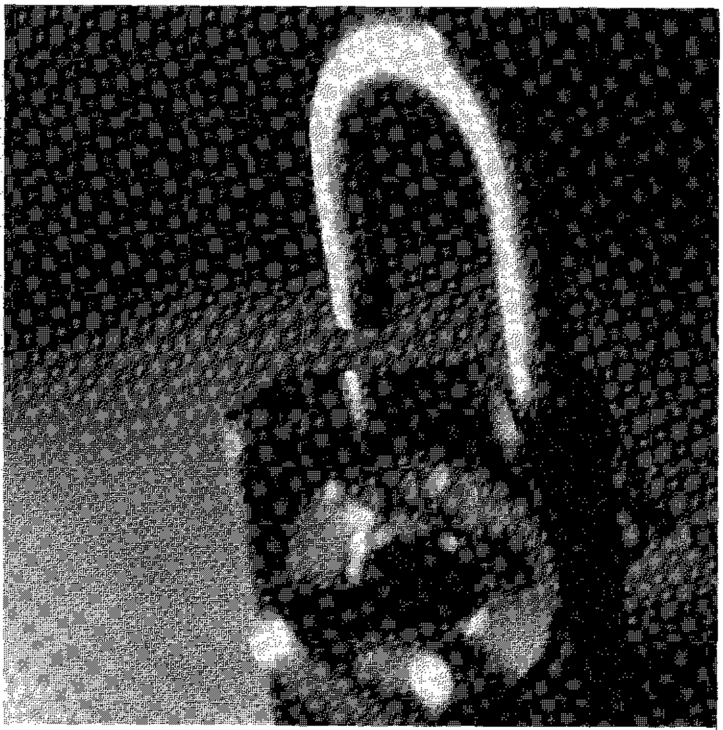
Not all school districts have been so resistant. Officials in Portland, Oregon, have gone out of their way to make parents aware of their opportunities. In October 2001, even before the federal bill had passed, the district sent letters to parents of students in three high schools it expected to land on the failing list, notifying them that the transfer option might be available for the next fall. During the summer of 2002, the district mailed a follow-up notice telling parents that the law provided transportation money for any students who wanted to leave those schools. Looking at test results in fall 2002, the district concluded that an additional middle school was likely to fall onto the failing list for the 2003–04 school year. Once again, it notified eligible parents and urged them to attend an annual school fair where they could learn more about the schools available for transfer.

Yet even with this ambitious effort—which far exceeds the outreach in the vast majority of cities—only about 140 parents sought transfers, says Lew Frederick, director of information at the Portland Public Schools. The reasons that so many parents decided to stay put, says Frederick, include: loyalty to neighborhood schools; reluctance to travel long distances; concern about how their children might be received in a new school; the already available opportunities to transfer under magnet programs; and a desire to gain access to the after-school tutoring services that are promised to students who remain in failing schools.

The record is clear that most districts could make it much easier, and more attractive, for parents to move their children from failing schools. But the factors suppressing participation in Portland—which appears to have made a good-faith effort to implement the program—suggest that there may be a natural limit to how many parents will move their children from one conventional public school to another. The experience in Dayton, and to some extent in Chicago, defines another limit on the current programs: in many places, the schools that students can transfer into may not be enough of an improvement on the schools they are leaving to make it worth the trouble.

Turf Battles

In the months ahead, the limitations of the choice provisions



established under the federal education law are likely to drive the debate in two different directions. From the left may come increasing demands that children in failing inner-city schools be guaranteed the right to transfer into neighboring suburban schools.

Faced with examples of suburban schools' refusing to take transfers, liberal education reformers are beginning to argue that suburban districts should be required to accept such children. In effect, they are using the logic of the education reform law to reopen one of the most divisive issues of the school desegregation era: whether largely white suburban districts should be required to accept black and Hispanic inner-city children. In the 1974 *Milliken v. Bradley* decision, the Supreme Court ruled that courts generally could not require busing across district lines to achieve racial balance; but the coming months may see more calls from liberals for moving kids across district lines to fulfill the promise of the federal law. "Choice within districts does not provide real opportunity to most of the students who need it," Goodwin Liu, a former education official in the Clinton administration, wrote recently. "The reality is that most high-performing public schools are located in the suburbs."

Conservatives take a different lesson from the disappointing results of the law's public school choice provisions. Many argue that the resistance from local public school bureaucracies shows that the only way to create genuine alternatives for children in weak schools is to provide them with private school vouchers. "There has never been more powerful evidence about the need for private school choice than the data that are coming out about public school choice," says Clint Bolick, vice president of the Institute for Justice, a conservative legal group. "In order to make the promise of No Child Left Behind meaningful, it's clear we have to look to every possible alternative, including private schools."

Some Bush administration officials are reaching the same conclusion. Eventually, President Bush may use the bureaucratic resistance to public school choice to revive the proposal for private school vouchers that he dropped early in the negotiations over the education bill in 2001. "That's definitely something we have given a lot of thought to," said one senior official.

In the meantime, the administration is trying to nudge districts toward more enthusiastic implementation of the law. In October 2002, the federal Department of Education distributed nearly $24 million in grants to Arkansas, Florida, Minnesota, and districts in six other states to expand their public school choice programs. More important, last fall the department also issued regulations announcing that districts could no longer use a lack of capacity as an excuse to deny transfers to students in failing schools. That alone will require many cities to intensify their efforts. "We know we are going to have to be much more aggressive next year," says Chicago's Hansen.

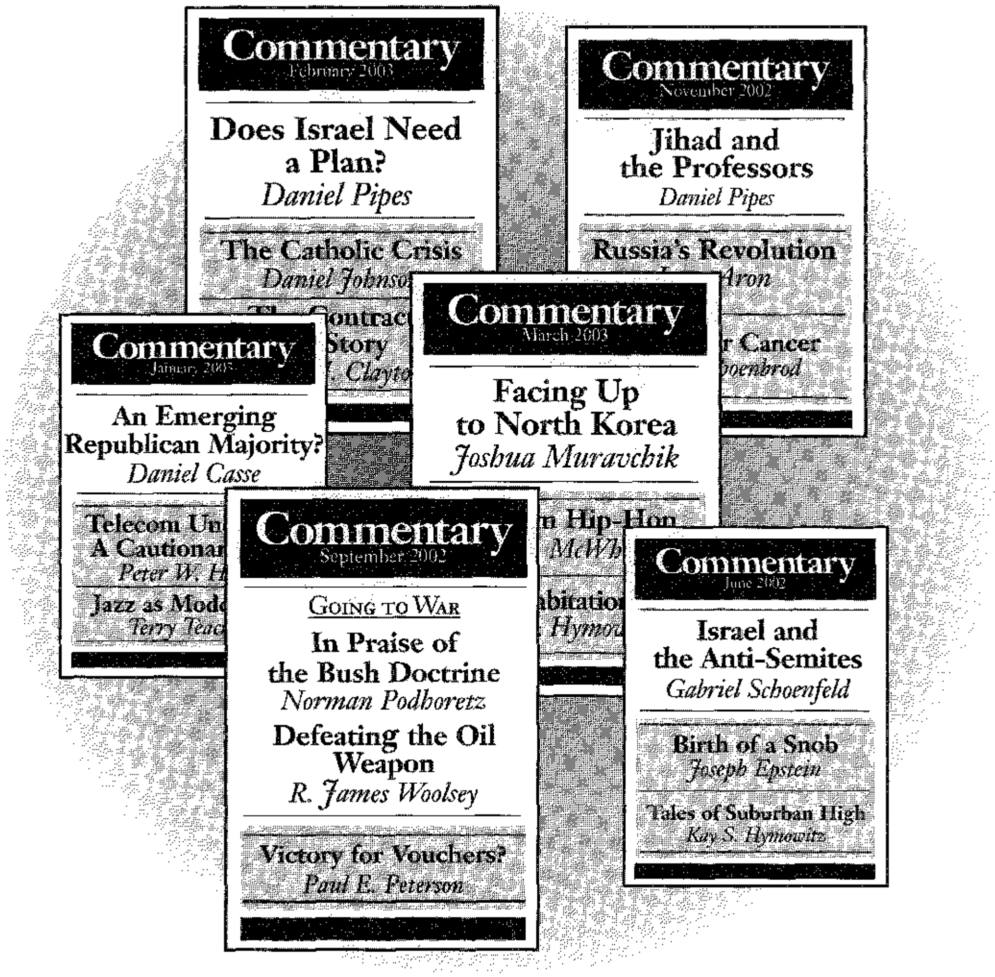
Whether that will be enough, in Chicago and elsewhere, to provide real opportunity for children trapped in failing schools remains very much in question. Liberals and conservatives may not agree on the cure, but both increasingly believe that the federal law's current approach to school choice is fatally flawed. The cool response to the transfer option in cities as different as Chicago and Portland suggests that, whatever choices are offered, many parents would rather see money and effort directed toward improving their neighborhood school. However, parents who want to vote with their feet (as Ronald Reagan once said) may need more opportunities than the federal reform law has provided so far.

The Bush administration may use the resistance to public school choice to revive a proposal for private school vouchers.

–Ronald Brownstein is a national political columnist for the Los Angeles Times.

"Very, very important to the life of the United States, to the West, and, I am convinced, to freedom."

—*Jeane J. Kirkpatrick*



Commentary
February 2003

Does Israel Need a Plan?
Daniel Pipes

The Catholic Crisis
Daniel Johnson

Commentary
November 2002

Jihad and the Professors
Daniel Pipes

Russia's Revolution

Commentary
January 2003

An Emerging Republican Majority?
Daniel Casse

Commentary
March 2003

Facing Up to North Korea
Joshua Muravchik

Commentary
September 2002

GOING TO WAR

In Praise of the Bush Doctrine
Norman Podhoretz

Defeating the Oil Weapon
R. James Woolsey

Victory for Vouchers?
Paul E. Peterson

Commentary
June 2002

Israel and the Anti-Semites
Gabriel Schoenfeld

Birth of a Snob
Joseph Epstein

Tales of Suburban High
Kay S. Hymowitz

Commentary

Sound thinking for the 21st century.®

$45.00 for 11 issues • Subscribe today! • $79.00 for 22 issues

Call 1 800 829-6270

HIGH-STAKES research

The campaign against accountability has brought forth a tide of negative anecdotes and deeply flawed research. Solid analysis reveals a brighter picture

by MARGARET E. RAYMOND & ERIC A. HANUSHEK

"MAKE-OR-BREAK EXAMS GROW, BUT BIG Study Doubts Value" intoned a front-page *New York Times* headline in December 2002. The article continued, "Rigorous testing that decides whether students graduate, teachers win bonuses, and schools are shuttered ... does little to improve achievement and may actually worsen academic performance and dropout rates, according to the largest study ever on the issue." Thus a deeply flawed study was catapulted to national prominence. More important, its conclusions were opposite those found through rigorous scientific studies.

The report in question, authored by Arizona State University researchers Audrey Amrein and David Berliner, purported to examine student-performance trends on national exams in states where legislators have attached "high stakes" to test scores. High-stakes testing has become a lightning rod as more and more states adopt accountability measures in response to the mandates of the federal No Child Left Behind Act. While it is crucial to analyze and debate the wisdom of such poli-

ILLUSTRATION BY ANDREW JUDD/MASTERFILE

cies, the discussion must be informed by evidence of the highest quality. The controversial nature of high-stakes testing has led to the hurried release and dissemination of research that lacks scientific rigor, of which the Amrein and Berliner study is one of the more egregious examples.

This says much about the standards for research in education today. The situation is so contentious that in 2000 the National Research Council found it necessary to convene a panel to decide which scientific principles should apply to educational research—the kind of question that other fields of social science settled long ago. In the case at hand, Amrein and Berliner trumpet the fact that their report was reviewed by a panel of four scholars based at other schools of education, yet this should only be a source of greater concern. Sharing a paper with sympathetic colleagues is no substitute for a system of blind peer review—a bedrock principle of scientific research.

Amrein and Berliner's misleading reporting practices took on new importance when the media dutifully broadcast the results as they were written.

Here we closely examine Amrein and Berliner's underlying data and methodology. Our results are astonishing: if basic statistical techniques are applied to their data, it reverses nearly every one of their conclusions. Later we also present the results of separate research on accountability that we conducted for a June 2002 Federal Reserve Bank of Boston conference. Rigorous analysis reveals that accountability policies have had a positive impact on test scores during the past decade.

The Unscientific Method

Amrein and Berliner identified 28 states where test scores are used to determine various consequences, such as bonuses for teachers, the promotion of students, or allowing children to transfer out of a failing school. These stakes go beyond less controversial accountability measures such as publishing test scores in the newspaper. The states range from Georgia and Minnesota—where the only penalty is experienced by students who fail a high-school graduation exam—to North Carolina and Texas, where the authors found a total of six stakes each, stakes that affect both schools and students.

Once Amrein and Berliner identified the high-stakes states, they looked at changes in the average scores students earned on the National Assessment of Educational Progress (NAEP). Choosing this test as a basis for considering the impact of high-stakes tests on students in the 4th and 8th grades (ages 9 and 13, respectively) is a sensible idea, because the validity and reliability of NAEP, often called the "nation's report card," are well accepted. It is a test for which students cannot easily be prepped and, since the performance of individual school districts, schools, or students is not reported, there is little incentive to cheat or even to prepare for the test. It also provides a neutral standard for assessing the effects of state policies. But if the Arizona State team's decision to look at NAEP scores was correct, less can be said for their other analytical choices.

Amrein and Berliner's basic strategy was to look at how each high-stakes state's scores changed with the introduction of accountability and to compare this with the national trend. If the state's gains exceeded the national gains, they deemed that an increase in scores. If the state's gains trailed the national gains, they deemed that a decrease. But whenever the rate at which students were excluded from the NAEP because of a disability or lack of language proficiency moved in the same direction as that state's NAEP scores (in other words, an increase in test scores coupled with an increase in test exclusions), Amrein and Berliner declared the results contaminated and simply tossed out the state as inconclusive. (At least that is what they claimed to do; in fact, they applied the rule inconsistently.)

As a result, their conclusions are based on only a fraction of the high-stakes states. For instance, they recorded positive or negative results on the NAEP 4th-grade math test for just 12 of the 26 states with stakes for K–8 students (as noted earlier, two of the states, Georgia and Minnesota had only a high-school graduation exam and thus were not used for this analysis). Amrein and Berliner found that 4th-grade math scores increased at a slower rate than the national average in 8 of the 12 states, faster in just 4. Yet they write this up in a highly misleading fashion, claiming that "67 percent of the states posted overall decreases in NAEP

Rerunning the Amrein-Berliner Data (Table 1)

When the actual test scores in the states Audrey Amrein and David Berliner identified as "high stakes" are compared with those in states without accountability systems, the high-stakes states show much more improvement.

	Increase in NAEP 4th-grade math scores		Increase in NAEP 8th-grade math scores	
	1992-2000	1996-2000	1992-2000	1996-2000
High-stakes states	9.2	4.2	8.8	4.5
No accountability states	3.8	2.3	4.0	1.7
High-stakes advantage	**5.3 points***	**1.9 points***	**4.8 points***	**2.8 points***
High-stakes advantage after adjusting for changes in students excluded from NAEP	5.2 points*	2.3 points*	3.7 points*	2.5 points*

* statistically significant at the .05 level

SOURCE: Authors

math grade 4 performance as compared to the nation after high-stakes tests were implemented." Actually, Amrein and Berliner witnessed gains slower than the national average in just 8 of 26 high-stakes states, or 31 percent.

Amrein and Berliner's misleading reporting practices took on new importance when the media dutifully broadcast their results as they were written. Consider the article in *Education Week*, which reported, "Movement in elementary-school reading scores was evenly split—better than the national average in half the states, worse in the other half." In fact, Berliner and Amrein based their conclusions in 4th-grade reading on just ten states, five of which they recorded as gaining against the national average, five of which as losing. So less than a fifth of the high-stakes states saw decreases against the national average in reading, not "half." At the 8th-grade level in math, Amrein and Berliner were able to look at only eight states, five of which gained against the national average, three of which lost. Here, again, Amrein and Berliner wrongly reported this as "63 percent of the states posted increases in NAEP math grade 8 performance as compared to the nation after high-stakes tests were implemented."

All of this ignores the truly fatal flaw of Amrein and Berliner's methods: their point of comparison. If one wants to assess the effect of high-stakes testing, the obvious comparison is between states that adopted accountability systems and those that did not. Amrein and Berliner's decision instead to compare the gains in high-stakes states with the national average violates a basic principle of social-science research. The national gain on NAEP incorporates any gains in high-stakes states, so Amrein and Berliner's strategy is akin to a medical trial where the treatment group receives the full dose of a medication while the control group receives a half-dose. It would not be surprising to find that the full dose was not dramatically more effective. The real question is whether the full dose is more effective than no medication at all.

On Their Terms

Amrein and Berliner concluded, as announced in their press release, "High-stakes tests may inhibit the academic achievement of students, not foster their academic growth." Let's take a look at their evidence in more detail.

Before doing so, however, we need to be clear: we are not in any way endorsing Amrein and Berliner's analytical approach. We return below to discuss the results from a more scientific study of accountability. But using their approach in a systematic manner will at least reveal the degree to which their decisions about what information to include and to exclude distorted the facts and thereby confused the debate over accountability.

An initial problem with their analysis is that Amrein and Berliner disregarded the magnitude of any changes in test scores. By simply listing the results as "Increase," "Decrease," or "Unclear" (in cases where exclusion rates rose), Amrein and Berliner discarded rich information. They converted useful continuous data (test scores) into hollow binary data (test scores went up or down). In a purely hypothetical example, say six of the high-stakes states gained 20 percent, while the other 20 gained 2 percent each and the no-accountability states made no gains whatsoever—yielding a national average gain of 3 percent. Amrein and Berliner's approach would supposedly demonstrate the failure of accountability: just six states beat the national average, while 20 were below the

A Panoply of Mistakes (Table 2)

Test scores actually increased at a faster rate than in no-accountability states in almost all of the high-stakes states where Audrey Amrein and David Berliner (AB) claimed to find decreases in scores. In New Mexico, Oklahoma, and West Virginia, where AB found decreases, high-stakes testing was introduced too early to make a valid before-and-after comparison.

States where AB declared decreases in NAEP scores	Introduction of high-stakes testing (AB date)	Change in 4th-grade NAEP math scores between:		
		1992-1996	1996-2000	1992-2000
Kentucky	1994	4.9[b]	1.0	5.9[c]
Maryland	1993	3.4[b]	1.6	5.0[c]
Missouri	1993	2.5[c]	3.8[b]	6.3[c]
Nevada	1998	N/A	2.7[c]	N/A
New Mexico	1989[a]	0.5	0.0	0.6
New York	1999	4.2[b]	3.9[b]	8.1[b]
Oklahoma	1989[a]	N/A	N/A	4.7[c]
West Virginia	1989[a]	8.1[b]	1.5	9.6[b]

Notes:

N/A - NAEP data unavailable for this time period

a. No NAEP tests at or before introduction of high-stakes testing

b. Change in NAEP scores exceeds the average change in NAEP both for the nation and for states not adopting high-stakes testing

c. Change in NAEP scores exceeds the average change for states not adopting high-stakes testing

SOURCE: Authors

average. In fact, ignoring any complications from test exclusions, Amrein and Berliner would report this as something like, "Just 23 percent of states posted gains on NAEP higher than the national average after high stakes were introduced." The right approach is to compare the average gains of high-stakes states with those of no-accountability states.

When this is done, the analysis yields starkly different results than Amrein and Berliner report. Table 1 compares the math gains among 4th and 8th graders in the same way as Amrein and Berliner—by following different cohorts as they reach 4th or 8th grade in different years. In other words, they compared the 4th graders of 1996 with the 4th graders of 2000, two completely different cohorts of students. For each of the comparisons, data were available for 34–36 states, 18–20 of which were part of the high-stakes group, due to the varying participation of states in the NAEP testing program. For either the 1992–2000 period or the 1996–2000 period, the average gain in math among high-stakes states noticeably exceeded that of the no-accountability states. The differences in performance were statistically significant at conventional levels, meaning that we can be highly confident that they are not just chance occurrences. (By contrast, Amrein and Berliner did no significance testing whatsoever, neglecting one of the oldest and most basic tools of social-science research.)

Amrein and Berliner might object that we have included states where students were excluded from tests at higher rates after accountability reforms were introduced, possibly contaminating the results. Amrein and Berliner's solution was just to toss these states out, no matter how small the change in exclusion rate or how large the change in achievement. As Table 1 shows, we instead adjusted the achievement gains for observed changes in exclusion rates. And the results barely changed: high-stakes states still significantly outperformed no-accountability states across the board. In fact, the changes in test-participation rates were not statistically different in high-stakes states from those in other states, indicating that this was not even remotely as influential a factor as Amrein and Berliner declared it to be.

Scientific quality is determined not only by the overall methodology, but also by the care and precision of any measurements. To assess the latter, let's focus on the eight states where Amrein and Berliner concluded that 4th-grade math scores decreased following the introduction of high-stakes testing. Consider Table 2. Three of the eight states—New Mexico, Oklahoma, and West Virginia—adopted high-stakes testing during the 1980s. However, NAEP scores at the state level became available only during the 1990s. For these states, Amrein and Berliner lacked

Correctly applying Amrein and Berliner's underlying approach to all of the data on NAEP achievement reverses their conclusions.

the "before" data for their "before and after" analytical strategy, but went ahead and labeled their scores as "decreasing" anyway. The other five "decreasing" states all experienced greater gains than no-accountability states during the time that they introduced high-stakes testing; New York even beat the national average gain in every time period. And this is the group of states that Amrein and Berliner identify as being harmed by accountability! *Not a single one* provides evidence of harm following the introduction of high-stakes testing.

Even where before-and-after data were available, Amrein and Berliner did not always use the data from the NAEP tests immediately preceding and following the adoption of high stakes. In several cases, they apparently chose an interval that began after the state's accountability system came on-line—an "after-after" comparison. These procedures yielded results that reflected negatively on accountability, but they have no scientific justification. To see this, consider the table on p. 52 and try to think of a consistent rule that justifies Amrein and Berliner's decision to place both Maryland and Missouri in the "decreasing" category.

In short, Amrein and Berliner used scientifically inappropriate methods and applied them in an even shoddier manner. Simply taking Amrein and Berliner's approach and applying it correctly to all of the data on NAEP achievement reverses their conclusions. Again, these simple comparisons are not the best way to examine these questions, but the results of even these crude analyses confirm the findings from the more sophisticated approach we describe below: greater accountability is accompanied by improved student performance.

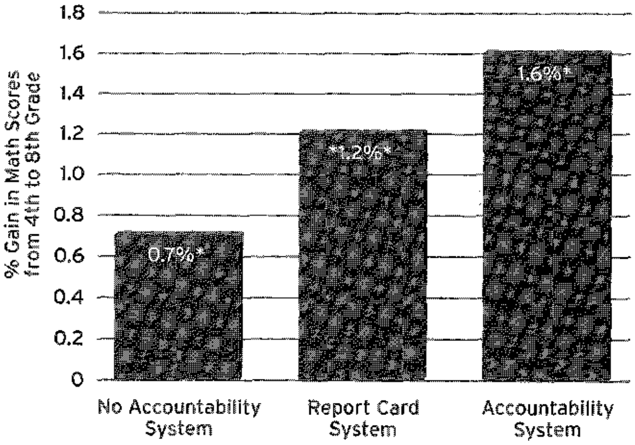
Amrein and Berliner also used trends on the SAT, the ACT, and Advanced Placement (AP) exams to assess the effectiveness of minimum-competency exams in the 18 states where students must pass such tests in order to graduate from high school. This comparison suffers from all the same problems as the NAEP comparison and more. For example, does anyone believe that nothing else has changed in North Carolina since the introduction of a graduation test in 1980? Amrein and Berliner's simplistic trend analysis attributes all subsequent changes in graduation rates and dropouts to the introduction of this high-stakes exam. Nonetheless, because these discussions are less directly related to the current state accountability debates and these data are more difficult to interpret than NAEP scores, we do not pursue them.

Results of Rigorous Analysis

Assessing the impact of state accountability systems is clearly complicated. In many states, these systems are quite young; in 1996, just ten states had active accountability systems. Moreover, states differ in many ways other than their accountability provisions—ways that can make it difficult to isolate the impact of high-stakes testing. They also change in different ways over time, adopting new accountability provisions and other legislation at different times and being influenced by shifting demographics at different rates. This does

Accountability Works (Figure 1)

States that reward or sanction schools for their academic performance made greater gains on the National Assessment of Educational Progress in math from 1996 to 2000.



*Statistically significant at the .05 level

SOURCE: Authors

States with high-stakes and even low-stakes accountability systems for schools performed significantly better on NAEP than states with no stakes at all.

not make gathering evidence about the effects of accountability impossible. It simply reinforces the need to apply stringent scientific methods to the analysis.

Here we report results from our own analysis of state accountability systems using NAEP data. These results were reviewed at a high-profile conference and were subject to a blind peer review for publication in a Brookings Institution volume, *No Child Left Behind? The Politics and Practice of Accountability*, which is slated for release this fall.

NAEP tested 4th graders in mathematics in 1992 and 1996 and 8th graders four years after each of these assessments, in 1996 and 2000. As noted earlier, whereas Amrein and Berliner simply compared the test scores of 4th graders in one year with those of a different set of 4th graders four years later, we measured students' growth in achievement between the 4th and 8th grades. In other words, we compared 4th graders' math achievement in 1996 with their performance four years later, when they were 8th graders. The same exact students were not tested in each grade, but the two samples are at least representative of the same cohort of students. We also adjusted the data to account for changes in state spending on education and for parents' educational levels, which provides controls for simultaneous changes in state policies or differences in demographics that might confound the analysis of how accountability systems influenced student achievement. Amrein and Berliner used no statistical controls at all.

Our analysis focuses on state testing and accountability systems that impose consequences on schools rather than on students. These are the most relevant policies for evaluating the potential impact of the No Child Left Behind Act. Our statistical analysis includes all states that have relevant NAEP data, and we explicitly allow for the timing of states' introduction of their accountability systems.

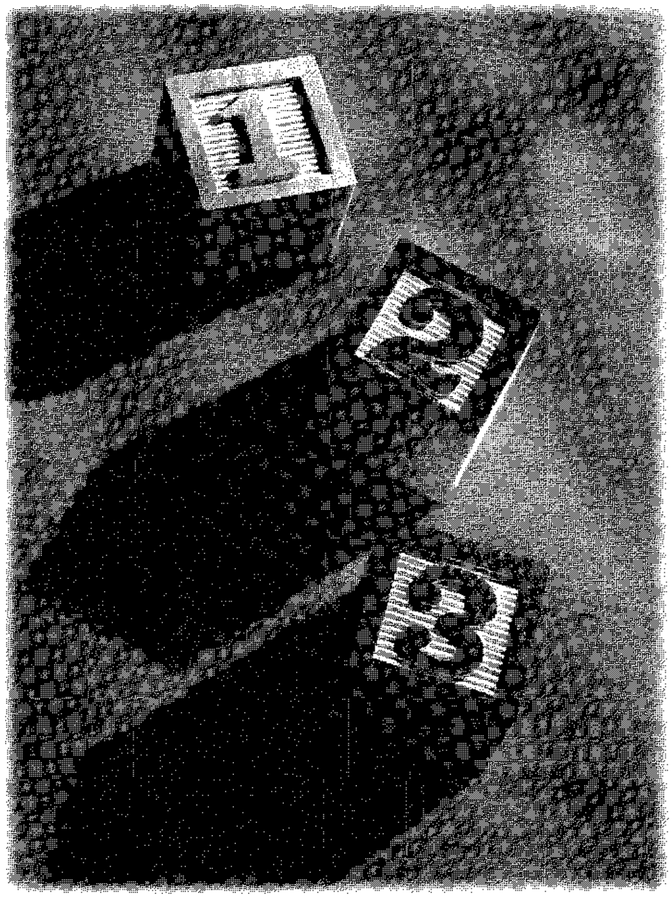


Figure 1 summarizes our findings in mathematics. The typical student progressing from grade 4 in 1996 to grade 8 in 2000 in a state with a consequential accountability system could expect to see a 1.6 percent increase in his NAEP proficiency score (calibrated to the appropriate learning standards for each grade). By contrast, the typical student in a state with no accountability system could expect to experience only a 0.7 percent gain in mathematics proficiency, a statistically significant difference. Students in states with "report card" systems, where scores are publicly reported but no consequences are attached to performance, fell in the middle: they could expect to gain 1.2 percent in achievement between grades 4 and 8, over and above what they would normally learn from grade to grade. In short, states with high-stakes and even low-stakes systems for schools performed significantly better on NAEP than states with no stakes at all.

We are not the only ones reporting positive effects of accountability. In a forthcoming paper, Stanford University economists Martin Carnoy and Susanna Loeb conducted a similar analysis but expanded it to include testing policies that impose high stakes on students. They found that NAEP performance increased in states with high-stakes systems compared with states that had not yet attached consequences to schools' test scores. Carnoy and Loeb also investigated the impact of accountability on student retention and high-school graduation rates and demonstrated that there is no discernible negative effect on either outcome.

Other rigorous studies have been carried out of accountability systems within states and school districts. As opposed to the Amrein-Berliner study, they have been vetted at scientific conferences and are being peer reviewed according to normal scientific practice. The Brookings Institution volume is one example. The accumulated literature generally supports two conclusions. First, student performance on the available measures, usually state tests, improves after accountability reforms are introduced. Second, other short-run changes—such as students' being excluded from taking the tests at greater rates, or explicit cheating—are observed. In other words, some unintended consequences often tend to accompany the introduction of accountability, although there is little evidence that these influences continue over time.

Schools may exclude low-performing students from taking the test in an attempt to "game" the system—to increase their performance artificially by removing scores that bring down their averages. We looked at differences among the states in terms of their placement rates into special education—often one way to exclude students from state tests—and at whether these differences were related to the introduction of state accountability systems. From 1995 to 2000, the time when many state accountability systems were coming on-line, we found no evidence that special-education placement increased in reaction to the introduction of accountability. Special-education placements did increase nationally, just not in any systematic way suggestive of a relationship to state accountability.

No Accountability for Research

That a study of such dubious scientific quality could make the front page of the nation's most respected newspaper is disturbing, but perhaps not so unusual. In the contentious environment of K–12 education, the media too often gives attention to findings that are relevant to policy regardless of their scientific merit. This discussion shows that education studies vary so much in their scientific rigor that one cannot just review them based on press releases and the sensationalism of the reported results.

Reporters need not be experts in statistical analysis any more than they must be fully versed in biochemistry or investment-banking regulations. But when a report is commissioned by an organization like the Great Lakes Center for Education Research and Practice, a Midwestern group sponsored by six state affiliates of the National Education Association, it would seem to call for a reasonable dose of skepticism. Why not bring in some outside expertise to review such a report before heralding its arrival? There will definitely be further opportunities for review. After all, the Arizona State shop promises that this is just the first of many annual reports on the impact of high-stakes testing.

The media is not alone. Resources at the state and federal levels must be committed to evaluating the quality of research and disseminating evidence of effective practices to schools and the public. The No Child Left Behind Act's emphasis on research-based practices, the creation of the federal Institute of Education Sciences, and efforts such as the What Works Clearinghouse, which will review and disseminate research findings, are important developments in this regard. State policymakers must also devote resources to evaluating their programs and synthesizing available research. Identifying effective reforms using rigorous evaluative techniques is a crucial task, especially since improving the education system is likely to have a greater economic impact than any of the medical breakthroughs of the past decade.

We also do not mean to suggest that the book has been closed on accountability. It appears that high-stakes states performed better than no-accountability states during the 1990s, but there is still much to be learned. For instance, there is uncertainty about the best way to translate test scores into overall school ratings. Also, states have yet to design accountability systems that directly link test-score performance to appropriate incentives. The vast majority of state accountability systems simply report the average scores for each school, sometimes disaggregating by racial and ethnic groups. However, average scores are highly dependent on socioeconomic factors outside the control of schools. States—and researchers—must become adept at discerning the components that make up the scores and how they can be influenced by high-stakes regimes. Measuring the gains that students make over time would provide a better measure of school performance and serve as a proper basis for reward or sanction, but such value-added techniques need some work before they can serve as reliable performance measures. There are other issues as well. Nonetheless, the evidence points in the direction of refining accountability systems rather than scrapping them altogether.

–Margaret E. Raymond is the director of CREDO, an education policy research group at the Hoover Institution. Eric A. Hanushek is a senior fellow at the Hoover Institution.

CROWD

An international look at the relationship between class siz

REDUCING CLASS SIZES IS ONE OF TODAY'S MOST popular education reform strategies. The Education Commission of the States estimates that such efforts cost states $2.3 billion during the 1999–00 school year alone. The federal government contributed another $1.6 billion in 2000–01 toward meeting the Clinton administration's goal of decreasing class size nationwide in the early grades to no more than 18 students. During the past year or so, the deteriorating condition of state budgets and the Bush administration's new emphasis on accountability have made class-size reduction less of a priority. Yet it remains popular among parents, teachers, and the teacher unions, which often promote it as an alternative to vouchers.

The motivation for reducing class size is intuitive: with smaller classes, teachers should be able to devote more time to each student, both in the classroom and in giving feedback on homework and tests. The concern is at least threefold. First, reducing class size is remarkably expensive, since it requires hiring more personnel. There may be less costly reforms that are at least as effective as class-size reduction. Second, hiring more teachers may dilute the quality of the workforce, thereby negating any gains among the students of good teachers.

PHOTOGRAPHS BY GETTY IMAGES

Control

and student achievement

by MARTIN R. WEST & LUDGER WOESSMANN

Finally, the intuitive relationship between class size and teachers' effectiveness may not actually hold true—teachers may be no more successful with 18 students than with 23.

The most persuasive evidence of the benefits of class-size reduction has come from the Project STAR (Student/Teacher Achievement Ratio) experiment in Tennessee, where students were randomly assigned to classrooms of varying size. Smaller classes appeared to yield substantial gains among kindergartners and possibly 1st graders in the first year of the program—gains that were maintained throughout their school years. However, a large body of research literature on class-size reduction contradicts the findings from Project STAR.

To lend a fresh perspective on this issue, we use data from the Third International Mathematics and Science Study (TIMSS) to compare the effects of class size around the world. While Americans squabble over whether class size should be 18 or 25 students, teachers in Korean schools routinely face classrooms of more than 50 students. These and other differences, such as the quality of a nation's teachers, can be valuable tools in discerning where, if ever, class-size reductions are likely to be beneficial.

Two Strategies

Ascertaining the effect of class size is less straightforward than it might appear. The central problem is that students are not assigned to classrooms randomly. For instance, schools often establish small remedial classes for lagging students or small enrichment classes for the so-called gifted and talented. In addition, school systems may direct students into schools with different average class sizes on the basis of their performance.

Parents also may influence their children's class sizes. They may work hard to move their children to schools with smaller classes, where they are likely to receive more attention. Thus variation in class size may be simply the result, rather than the cause, of differences in student achievement. Estimating the true effect of class size on student performance requires a strategy that looks only at variations in class size that are unrelated to students' previous achievement.

In principle, two such strategies are available. The first is to conduct a randomized field trial along the lines of Project STAR in Tennessee. Unfortunately, while it used a powerful research design, the Tennessee study was flawed in its implementation. For one thing, no data were collected on students' performance before they were assigned to their classrooms, making it impossible to know whether the assignment was truly random. In addition, the teachers were aware of their participation in Project STAR, as in almost any true experiment. This has led some to question whether its findings can be expected to hold under more typical conditions. It is also worth noting that the evidence here comes from an experiment conducted in a single U.S. state during the mid-1980s, in which classes were reduced from 22–25 students to fewer than 17. In that sense, the findings may not apply to school systems in other parts of the world.

The second strategy, quasi-experimental research, relies either on special types of variation in class size or on econometric techniques to make appropriate comparisons. However, the conditions that must be met in order to use this approach make credible quasi-experimental studies possible for only a small number of school systems. For example, Princeton economists Anne Case and Angus Deaton used data on black students in South Africa during apartheid to measure the effects of class size. They argued that the black population of South Africa during this time lacked the power to influence class sizes, making the assumption that students were randomly assigned to classrooms of different size more plausible. But the South African school system under apartheid was obviously unique; in some districts, the average class size reached 80 students.

In just two countries, Greece and Iceland, did smaller classes appear to elicit superior student performance.

While Case and Deaton found that smaller classes were modestly beneficial, Harvard economist Caroline Hoxby's careful quasi-experimental study of elementary schools in Connecticut suggests that Case and Deaton's results may not be relevant for more developed countries. Hoxby analyzed variation in class size due to random fluctuations in the number of births and restrictions on maximum class sizes. She found no evidence of even trivial class-size effects. However, her approach requires a long panel of rich data and has yet to be applied in other contexts.

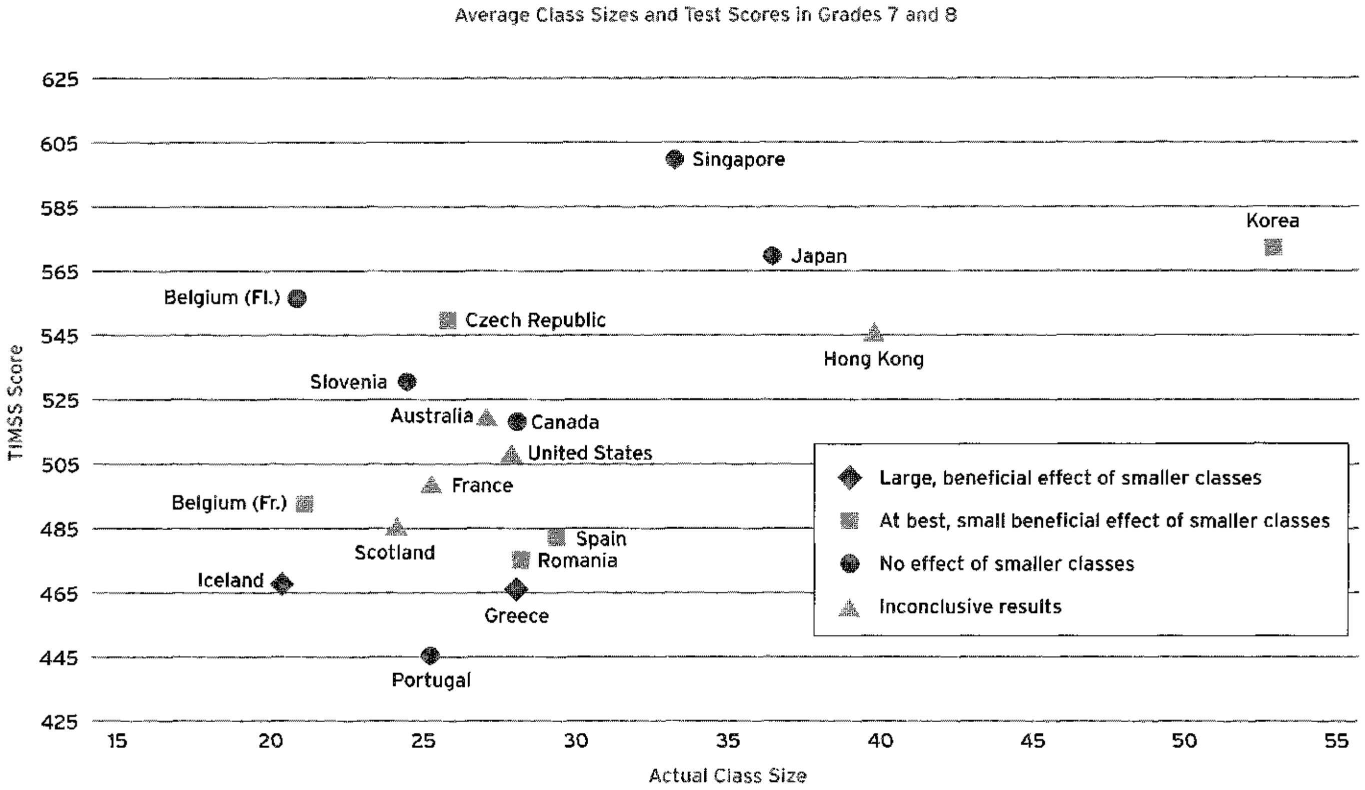
International Evidence

Taking data from TIMSS, we used a quasi-experimental design to take a broader look at how class size affects student achievement in different nations around the world. Conducted in 1994–95, TIMSS was the largest international study of student performance ever, with more than 40 countries participating initially. Each country administered the test to a nationally representative sample of middle-school students, defined as those students enrolled in the two adjacent grades that contained the largest proportion of 13-year-old students at the time of testing (grades 7 and 8 in most countries).

Our strategy takes advantage of the fact that data were collected on both actual and average class sizes and on students' performance and socioeconomic backgrounds for more than one grade level in each school. We looked at whether 7th

Few Benefits from Smaller Classes (Figure 1)

In just 2 of 18 countries, Iceland and Greece, were there large, beneficial, statistically significant effects of smaller classes. Large effects can be ruled out in 11 of the 18 nations.



SOURCE: TIMSS

graders in a particular school performed better than the same school's 8th graders (relative to the national average for their respective grades) when, on average, the 7th-grade classes were smaller than the 8th-grade classes. With this strategy, the variation in class size we considered is strictly a consequence of fluctuations in the cohort size from one grade to the next. This excludes variation in class sizes within the same grade and from school to school, both of which can be subject to the influence of parents and school-system policies that tend to sort students into classrooms by their performance. The remaining differences should be essentially unrelated to student performance.

This approach forced us to restrict the sample to schools in which both a 7th-grade and an 8th-grade class were actually tested and in which data on the actual class sizes and average class sizes were available for each grade. We ultimately conducted our analysis on the 18 countries in which data for at least 50 schools in both mathematics and science remained after applying these criteria.

As shown in Figure 1, Portugal exhibits the lowest average combined test scores in math and science among the 18 countries in our sample, Singapore the highest. Iceland has the smallest average class size, with just 20 students per classroom. At nearly 53 students per class, Korea has by far the highest average. The other East Asian countries also feature large classes, with an average of more than 30 students. In general, the countries with the smallest classes tended to be the worst performers. The reverse is also true: high performers tend to have larger classes. While this does not say much about the effectiveness of reducing class sizes in various environments, it does demonstrate that it is possible to have a high-achieving school system with relatively large classes.

Results

Let's look first at the results of a straightforward comparison that adjusts the data on student performance for students' socioeconomic background and grade level (since 7th and 8th graders were tested), thereby attempting to isolate the effects of class size. This initial analysis is of interest primarily because it is analogous to the approach used in most research on class size. Comparing these results with those obtained by a more reliable strategy will provide an indication of what biases may

East Asian countries feature large classes, with an average of more than 30 students.

exist in other studies.

In 11 of the 18 nations, the estimate of the effects of class size were positive and statistically significant, suggesting that students in larger classes perform significantly better than students in smaller classes. In other words, a naïve strategy that does not account for the ways in which students are sorted into classes of different size leads to the counterintuitive result that students fare better in larger classes. Moreover, this result seems universal: it emerges in western Europe (Belgium, France), eastern Europe (Czech Republic, Romania), Australia, and East Asia (Hong Kong, Japan). No country showed students in smaller classes outperforming their peers in larger classes.

Let's turn now to the preferred strategy, which controls for the fact that students performing at different levels may be sorted into smaller or larger classes both between and within schools. The first notable feature of this approach is the disappearance of the counterintuitive result that students do better in larger classes. In 16 of the 18 countries, none of the results was statistically different from zero. In the other two countries, Greece and Iceland, smaller classes did appear to elicit superior student performance. Moreover, the benefits appear to be substantial: Students scored just over two points (or 2 percent of the international standard deviation) higher for every one student fewer in their class.

This evidence suggests that capable teachers are able to promote student learning equally well regardless of class size.

Precision Testing

What can be learned from the 16 countries where the results were statistically insignificant? Does this suggest the lack of a causal relationship between class size and student performance? Or is it merely the result of statistical imprecision? In four of the countries, Australia, Hong Kong, Scotland, and the United States, the standard error of the estimated effects of class size was extremely large, indicating that little confidence should be placed in the results. The lack of precision in these cases seems to be a direct consequence of our research strategy's rather demanding data requirements. These school systems simply exhibit little variation in average class size from one grade to the next—the type of variation on which our strategy relied.

The remaining 12 countries can be further distinguished by comparing their results with those from other studies. We chose first to compare our results with those reported by Princeton economist Alan Krueger in his reanalysis of the Project STAR data from Tennessee, which produced some of the highest estimates of class-size effects among credible studies. Krueger performed a very rough cost-benefit analysis, in which the eco-

Low-Quality Teachers? (Table 1)

Iceland and Greece maintain both low class sizes and low school spending by paying their teachers poorly, suggesting that smaller classes may be beneficial only when teachers are of low quality.

	Countries*	Average Class Size	Average Score on TIMSS	Per-Pupil Spending	Average Teacher Salary	Teacher Salary/Per-Capita GDP	Share of Teachers with Each Level of Education				
							a.	b.	c.	d.	e.
Countries with large, beneficial effects of smaller classes	Greece Iceland	24	467	$2,374	$16,311	1.0	16%	5%	45%	32%	1%
Countries where smaller classes evince, at best, a small, beneficial effect	Belgium (French) Czech Republic Korea Romania Spain	31	514	$3,478	$27,496	1.8	2%	19%	15%	35%	27%
Countries where smaller classes have no effect	Belgium (Flemish) Canada Japan Portugal Singapore Slovenia	28	537	$5,667	$29,038	1.8	2%	32%	7%	55%	5%

*The results for Australia, France, Hong Kong, Scotland, and the United States were inconclusive

a. Training but No Secondary Education b. Secondary Education c. B.A. d. B.A. + Training e. M.A. + Training SOURCE: TIMSS and OECD

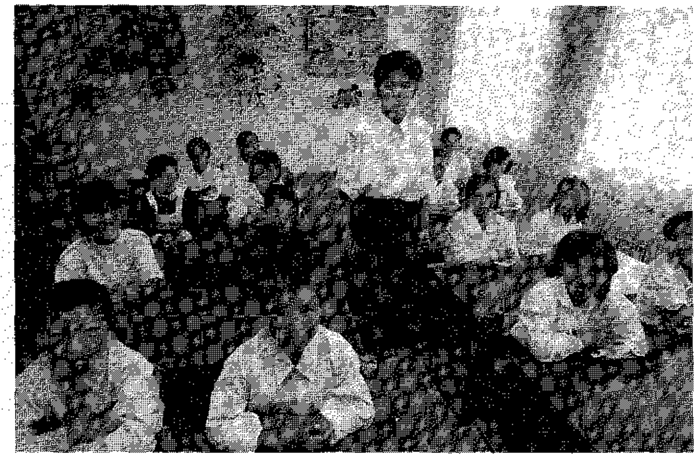
nomic benefits of class-size reduction, in terms of the increase in future earnings due to higher test scores, appeared to approximate the costs.

Krueger's results indicate that students in kindergarten classrooms that had 7 to 8 fewer students than regular-sized classes performed about 3 percent of a standard deviation better for every one student fewer in their class. Converted to international scores on TIMSS, this is equivalent to three test-score points. This is greater than the two-point gain we found in Iceland and Greece, but it is within the standard error of these estimates, suggesting that the actual effect of reducing class size in Iceland and Greece could be as large as Krueger found in the United States.

For 11 of the 12 countries with relatively precise yet statistically insignificant estimates, the possibility of class-size effects of the same size as Krueger found can be rejected with at least 95 percent confidence. There could still be class-size effects in these nations, just not of the magnitude estimated by Krueger. Note, however, that Krueger's effects were found in kindergarten and 1st grade, while these estimates are for students in 7th and 8th grades.

We further tested to see whether a one-student reduction in class sizes would increase TIMSS scores by just one point, or 1 percent of an international standard deviation. An effect of this size would be so small as to be essentially negligible from the standpoint of public policy; a one-point gain is too little to justify the expense of class-size reduction. Regardless, even the possibility of this small an impact can be rejected with at least 90 percent confidence in 6 of our 12 school systems with reasonably precise results.

In short, the effect of class size on student performance varies across the 18 countries in our sample (see Figure 1). We can rule out even a minimal relationship between class size and TIMSS scores in the middle grades in six school systems: those of Flemish Belgium, Canada, Japan, Portugal, Singapore, and Slovenia. In an additional five school systems, we can rule out the possibility of large class-size effects: French Bel-



The evidence suggests that capable teachers are able to promote student learning equally well regardless of class size.

gium, the Czech Republic, Korea, Romania, and Spain. These results cast doubt on the desirability of class-size reduction in the middle grades as a reform strategy in many countries. In Greece and Iceland, by contrast, smaller classes were clearly beneficial. (In five countries—Australia, France, Hong Kong, Scotland, and the United States—our strategy led to inconclusive estimates that do not allow for any confident assertions about the effects of differences in class size.)

Quantity versus Quality

Why would class-size reduction elicit improvement in Greece and Iceland but not elsewhere? One might expect class-size effects to be related to such characteristics as a nation's overall level of resources. For instance, it is feasible that countries with relatively large classes would glean substantial benefits from reducing class sizes. However, there is no clear pattern in countries' average class sizes that distinguishes the two countries where substantial class-size effects exist from either the six countries where we ruled out any noteworthy class-size effects or from the five countries where we ruled out at least large class-size effects. Greece's average class size is similar to the mean class size among the nations where no class-size effects were found, and Iceland's average class size is substantially lower (see Table 1).

One possibility is that class-size reduction has a large impact in relatively ineffective school systems. Both Greece and Iceland performed considerably below the international average on TIMSS, while the countries where class-size reduction did not have even a small effect performed above the average. Also, even though Greece's class sizes are roughly at the mean and Iceland's were substantially lower than the mean, education spending per student in both countries is substantially below the average of the two comparison groups. This suggests that Greece and Iceland spend rather little per employed teacher, which is reflected in the data on teachers' salaries. Teachers' salaries in Greece and Iceland are below the mean of the other countries in absolute terms, in terms of salary per teaching hour, and relative to the country's per capita GDP (see Table 1).

A low average salary for teachers suggests that a country may be drawing its teaching population from a pool of less-skilled workers. If this is the case, different countries appear to be making different tradeoffs between the quantity and quality of their teachers: with class sizes low, Greece and Iceland employ many teachers of low quality. The countries where class-size effects were not observed appear to employ relatively fewer teachers, but of higher quality.

This assumption is borne out by the available data on teachers' educational attainment. In Greece, the highest level of education reached by the vast majority of teachers is the equivalent of a bachelor's degree without any teacher training. In Iceland, about one-third of the teachers surveyed by TIMSS had not even completed secondary education, with only some basic teacher training. Meanwhile, about 60 percent of the teachers surveyed in the other countries held either a bachelor's or a master's degree in addition to their training as teachers.

This evidence suggests that capable teachers are able to promote student learning equally well regardless of class size (at least within the range of variation that occurs naturally among grades). Less capable teachers, however, do not seem to be up to the job of teaching large classes.

This interpretation is corroborated by teachers' responses in TIMSS when they were asked to what extent their teaching was limited by a high student-to-teacher ratio in their classroom. In Greece and Iceland, 45 percent of teachers reported that their teaching was limited "a great deal" by a high student-to-teacher ratio. The comparable statistics averaged only 19 percent and 25 percent among countries where no class-size effects and no large class-size effects were found, respectively. This is despite the fact that average class sizes in Greece and Iceland were lower than in either comparison group.

In short, our evidence suggests that the existence of class-size effects is related to the quality of the teaching force. Smaller classes appear to be beneficial only in countries where average teacher quality is low. If teacher quality is a key input in education, this interpretation can explain why class-size effects exist in some countries but not in others and at the same time why the countries in our sample where we did find sizable class-size effects also exhibit poor overall performance. Greece and Iceland exhibit class-size effects and poor performance because they employ a population of relatively less capable teachers, while other countries exhibit no class-size effects but high overall performance because they employ good teachers. This suggests that it may be better policy to devote the limited resources available for education to employing more capable teachers rather than to reducing class sizes. The merits of this admittedly speculative conclusion are a promising topic for future research.

–Martin R. West is a research fellow at the Harvard University Program on Education Policy and Governance and the research editor of Education Next. *Ludger Woessmann is a senior researcher at the Ifo Institute for Economic Research in Munich, Germany.*

THE NATIONAL INTEREST

Economics as Philosophy
Michael Novak

Spring 2003

EMPIRE?

James Kurth on *Imperial Migrations*

☆ ☆ ☆

On *The National Security Strategy*
Philip Zelikow Jack Snyder

☆ ☆ ☆ ☆ ☆

Strategy at Arms
Stephen Peter Rosen Anna Simons

☆ ☆ ☆

Gideon Rose on Bacevich's *American Empire*

Josef Joffe on Continental Divides
Adam Posen on Germany's Path to Economic Perdition

Sociologies of Islam

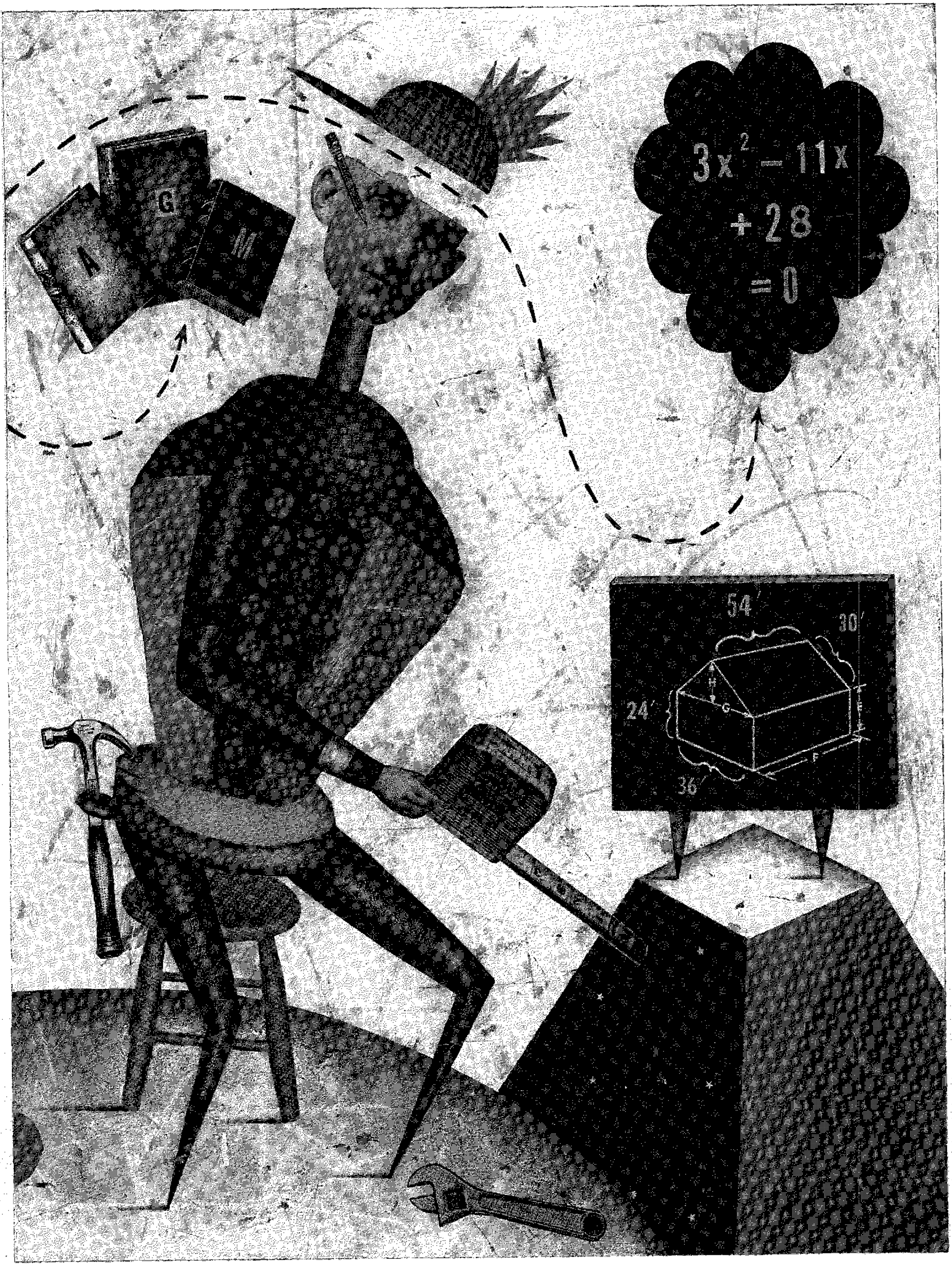
EuroIslam: The Jihad Within? *Olivier Roy*
Out of Bali: Cybercaliphate Rising *David Martin Jones*

Eisenhower on Khrushchev

Essays & Comments:
Michael Mandelbaum, Paul Hollander, Harvey Sicherman and others

Subscribe Now and Get 4 Issues for Only $26
Call (800) 893-8944 or (914) 962-6297

TPI0303


A
G
M
3x² − 11x + 28 = 0
54'
30'
24'
36'
H
G
E
F

learning to earn

More stringent high-school graduation requirements may reduce students' chances of earning a diploma. But higher standards also improve their ability to find a job

by THOMAS S. DEE

DURING THE 1970S, NEARLY EVERY STATE IN the nation began instituting tests of basic skills for high-school students as the leading edge of the so-called "first wave" of education reforms. These reforms were a response to the widespread impression that test scores and the quality of public schooling were in decline. According to critics, the high-school diploma, once a true accomplishment, had been debased in an era of social promotion and low standards—to the point where it held no real meaning for postsecondary institutions or potential employers.

The pace of reform was greatly accelerated with the release in 1983 of the blue-ribbon report *A Nation at Risk*. A chief cause of the nation's educational decline, the report ventured, was the "cafeteria-style curriculum" that allowed students to pursue a diffuse and unchallenging course of study. The report recommended that states require students to take a minimum number of courses in core academic subjects in order to graduate from high school. As a result, by 1992 nearly every state had increased its graduation requirements in the core academic areas. However, only three states, Florida, Louisiana, and Pennsylvania, had met

ILLUSTRATION BY COLIN JOHNSON

the standard recommended by the *Risk* report: four years of English and at least three years each of social studies, science, and math.

Such reforms, with their focus on testing and higher academic standards, are the precursors of today's controversial accountability movement. Yet there has been surprisingly little study of their consequences. For instance, did the requirement that all students pass a minimum-competency test in order to graduate from high school encourage more students (in particular, minorities) to drop out of high school, as many critics feared? Did such exams, as supporters hoped, make the high-school diploma more valuable, thereby improving the job prospects of graduates?

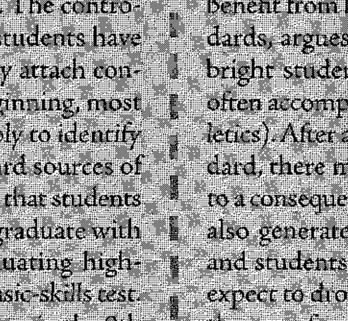
These questions continue to be relevant as the states ramp up their testing programs and impose ever-harsher sanctions on failing schools and low-performing students. Judging from their continuing prevalence in policy debates, such questions also remain mostly unanswered.

Controversies

Minimum-competency exams were by far the more controversial of the two major first-wave education reforms. In the end, few critics object to tougher course requirements. The controversy arises when states attempt to test whether students have met a certain standard—and especially when they attach consequences to the results of such exams. In the beginning, most minimum-competency exams were intended simply to identify low-performing students and to direct them toward sources of remediation. However, several states also mandated that students pass the minimum-competency exam in order to graduate with a standard high-school diploma. By 1992, graduating high-school seniors in 15 states were required to pass a basic-skills test.

Typically, students would first sit for these exams in the 9th or 10th grade and enjoy multiple opportunities for retests. The conventional wisdom has been that these tests were "legislated as a lion but implemented as a lamb." To wit, the exams were usually set at only the 8th- or 9th-grade level. The standards were sometimes set even lower in the face of politically unacceptable failure rates. As a consequence, the ultimate pass rates among high-school seniors were extremely high. The still-open question is whether these pass rates were artificially high, given that some students may have chosen simply to drop out if they thought they had little chance of passing the test.

Did requiring all students to pass a minimum-competency test in order to graduate from high school encourage more students to drop out?



Placing students at increased risk of dropping out is one of the major objections that critics lodge against any kind of testing regime that imposes harsh penalties on students. The intent of these reforms is both to raise academic standards and to give students the incentives to meet them. But critics say that standards-based reforms may simply exacerbate existing inequalities if sanctions are applied to low-performing students without giving them the resources and help they need to succeed. Furthermore, standards could lower the achievement of high-performing students if they signal that learning for its own sake is not worthwhile.

Economists Julian Betts and Robert Costrell have argued that students whose prior academic record provides a clear indication of how they will perform on the test (pass with flying colors or fail miserably) face little incentive to change their study habits. It is students at the margin who may feel most of the impact of a test- or curriculum-based sanction. Some will respond by redoubling their efforts to ensure that they pass the test or accumulate the proper number of credits in core academic subjects. Others may grow discouraged and eventually give up.

So the marginal students who choose to work harder may benefit from higher testing and curricular standards. High standards, argues Cornell economist John Bishop, may also shield bright students from the harassment and peer pressure that often accompany excelling in academics (as opposed to, say, athletics). After all, if all students must meet a defined external standard, there may not be much glory in failing and being subject to a consequence as harsh as not graduating. High standards may also generate broader educational gains as schools, teachers, and students struggle to attain them. Even students who fully expect to drop out may expend more effort in the short term as they are forced to progress through a state-mandated set of core academic courses. Thus both dropouts and graduates may reap the rewards of higher standards, if the standards signal to employers that the average ability of all students has increased.

To shed light on these debates, I examined national data on a variety of outcomes for students to assess the impact of the first wave of education reforms. The outcomes include both school-based achievements, such as students' course-taking and their level of educational attainment, and students' later experiences in the labor force.

Contradictions

The available evidence on the influence of minimum-competency exams and higher curricular

standards on educational attainment and employability is not only scant but often contradictory. One possible explanation for these conflicting results is that almost all previous empirical studies have relied on comparisons of students in states that adopted these reforms with students in states that did not. But test scores and other outcomes related to education vary dramatically from state to state, often for cultural, socioeconomic, and political reasons that are difficult for researchers to measure directly. In addition, these factors, which are unique to each state yet difficult to account for, may influence a state's decision to adopt reforms like minimum-competency tests and higher curricular standards. That makes it quite difficult to credibly isolate the effects of such reforms using traditional comparisons across the states.

Did such exams make the high-school diploma more valuable, thereby improving the job prospects of graduates?

To address these concerns, I used 1990 census data (the Public Use Micro-Data Sample), which allowed me to construct before-and-after comparisons of students from within the same states. I compared individuals of different ages, some of whom attended high school before the first-wave education reforms were implemented and others who attended afterward—and were thus required to take minimum-competency exams and more courses in academic areas in order to graduate. I made these comparisons by matching each state's graduating class to the policies that were in effect the year they graduated—whether students had to take a minimum-competency exam and whether they faced high course requirements in order to graduate. High course requirements were defined as a high-school curriculum that includes at least three years of English, two of social studies, one of science, and one of math. In the initial analysis, I assessed the impact of these reforms on the chance that students would graduate from high school and enter college. I then examined their effects on subsequent participation in the labor market and average weekly wages as adults.

In order to eliminate the influence of other national trends, I compared the changes within states that introduced first-wave reforms with the contemporaneous changes in states that did not. I also adjusted the data in both analyses to account for a variety of other within-state factors that changed over time and could have influenced student outcomes. For instance, first-wave reforms were sometimes part of omnibus education bills that included other policy changes, such as increased spending on K–12 education. To address this, I controlled for real per-pupil spending in the state when the survey respondents were 16 to 17 years old (in other words, respondents who were 18 years of age in 1980 were matched to their states' school spending during the 1978–79 school year). My results are also adjusted for the state unemployment rate when respondents were 17 years old, since high unemployment rates can provide an incentive for students to stay in school.

There is also evidence that educational attainment can be influenced by the relative size of a birth cohort. At the college level, this can occur if enrollment space at local colleges and universities is not expanded to accommodate temporary population booms. At the secondary level, larger student populations may increase class sizes and strain school resources, thereby lowering school quality and reducing the benefits of staying in school. For these reasons I introduced an adjustment for the size of students' birth cohorts. I also included a measure of the real costs of postsecondary tuition based on the in-state rate at "lower level" state colleges and universities when the respondents were 17 years old. Finally, as a control for within-state changes in socioeconomic conditions, I matched respondents to the poverty rates in their states when they were 17 years old.

One further note: I performed several tests to check whether my results were indeed more reliable than results using comparisons of students across states rather than within the same state. One method is to estimate the effect of a state-level policy that, in theory, should be totally irrelevant to educational outcomes. In this case I chose whether a state had any executions when the student was 18 years of age. In the across-state comparison, capital punishment generated large and statistically significant reductions in the probability of high-school completion and college entrance, clearly an implausible result. By contrast, the effects of capital punishment as estimated in the within-state comparisons were negligible, as they should be. This strongly suggests the increased reliability associated with using within-state comparisons.

Education and Employability

For all students, the minimum-competency exams showed no statistically significant effect on the probability of graduating from high school. By contrast, higher curricular standards reduced students' probability of graduating from high school by 0.5 percentage points (see Figure 1). An effect of this size represents a 3 percent increase in the probability of dropping out. Another way to frame the size of this estimate is to note that high-school completion rates among 18- to 24-year-olds increased from 82.8 percent in 1972 to 86.5 percent in 2000. My results suggest that in states that adopted high curricular standards these average gains were 14 percent less than they otherwise would have been. (Note that these percentages—as well as the data set I used—considered students who passed the General Educational Development, or GED, test to be high-school graduates. Thus my findings may actually understate the true reform-induced reduction in high-school completion.)

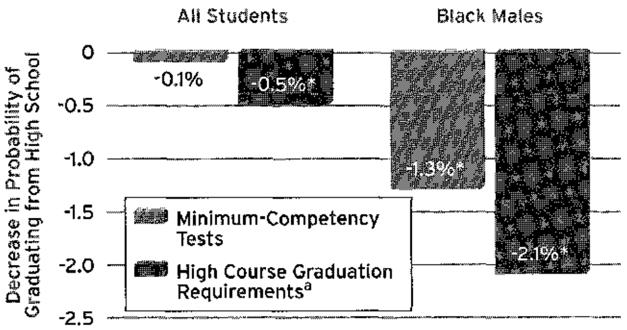
Higher course requirements significantly reduced the probability of graduating from high school for blacks and for white males.

For all students and for subgroups broken down by race and gender, the first-wave reforms had statistically insignificant effects on the probability of entering college. This makes sense since these reforms were not aimed at the relatively high-achieving students who are considering college.

However, in looking at the results for students separated by race and gender, both types of reforms had fairly large and *statistically significant effects on the probability of completing high school* for some groups. For instance, higher course requirements significantly reduced the probability of graduating from high school for blacks and for white males, but not for white females. Among black students, higher curricular requirements reduced the probability of graduating from high school by roughly 2 percentage points—four times the impact that these reforms had on white males. Similarly, the only *statistically significant effect of introducing a minimum-competency exam* was among black males, who experienced an estimated 1.3 percentage point drop in the probability of completing high school. These estimated effects are fairly large relative to the recent growth in educational attainment among blacks. Between 1972 and 2000, the high-school completion rate of 18- to 24-year-old black students increased from 72.1 percent to 83.7 percent. (Once again, these percentages include GED recipients.)

Effects of First-Wave Reforms (Figure 1)

Requiring students to pass minimum-competency exams and to take a heavier courseload[a] in order to receive a high-school diploma slightly reduced students' probability of graduating. The impact was greater among black males.



[a] High course requirements are defined as a high-school curriculum that includes at least three years of English, two of social studies, one of science, and one of math.

* *Statistically significant at the 5 percent level*

SOURCE: Author

These findings are largely consistent with the concerns sometimes raised by critics of standards-based reforms. The introduction of high-school graduation standards led to reductions in educational attainment that were particularly concentrated among minority students. But a full evaluation should also consider the implications of these reforms for students' performance in the labor market, especially since local business leaders concerned with the quality of their workforce were often instrumental in the adoption of first-wave reforms. As noted earlier, higher standards may benefit students (even those who drop out) by creating an incentive to work harder in school, an effort that is rewarded in the labor market. Higher standards may also increase the prestige of being a high-school graduate and correspondingly reduce the stigma associated with being a dropout.

My findings suggest that both first-wave education reforms had no statistically significant effects on adult wages for all groups. However, the effects on adult employment were more substantial (see Figure 2). Higher course requirements appeared to increase the probability of adult employment by roughly 1 percentage point for white males and by 3 percentage points for blacks, even after controlling for educational attainment. Furthermore, the introduction of minimum-competency exams increased the probability of adult employment among black males by 1.6 percentage points. One useful way to underscore the magnitude of these policy effects is to note that, in these data, white males were roughly 19 percentage points more likely to be employed than black males. Since the employment gains attributable to each first-wave reform were roughly 2 percentage points larger for black males than for white males, states that implemented them closed the black-white employment gap by roughly 11 percent.

Taken together, these results indicate that higher course requirements had educational and labor-market consequences for almost all students. The effects of minimum-competency exams, by contrast, were narrower, observable only among black males. These results are consistent with the anecdotal evidence suggesting that minimum-competency exams were often weakened in response to political realities. These results also suggest that, when binding, higher standards of either type had decidedly mixed consequences. They had negative effects on students' educational attainment, particularly among black students. But they also generated some employment rewards, again mainly for black students.

How can we compare these gains and losses? For those students who dropped out only in response to the new requirements, the loss was large: according to these data, high-school graduates can expect to earn approximately 33 percent more than the average dropout. This was only partially offset by improved prospects for employment. However, those students who would have dropped out regardless of the graduation requirements experienced significant gains because they were more likely to be employed.

Another way to frame the question is to ask how either type of reform might change the expected wage for students who were uncertain about whether they would be high-school graduates. My back-of-the-envelope calculations indicate that the expected wage benefits of higher curriculum standards due to their positive effect on employment exceed the expected wage costs due to their adverse effect on graduation rates by a factor of six. This suggests that a risk-neutral person might prefer a regime with higher curricular standards. However, a full cost-benefit analysis would need to account for not only the labor-market consequences but also the disparate impacts on various ethnic and socioeconomic groups and other social losses that might accompany increased dropout rates.

Standards on the Ground

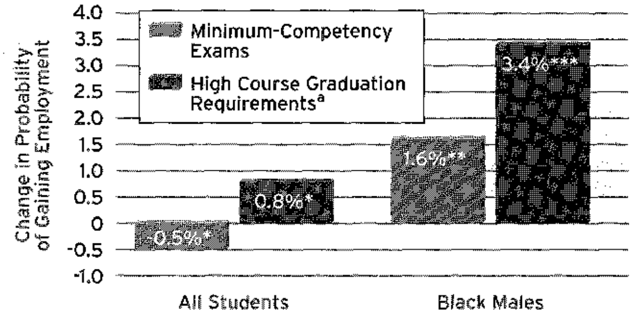
How did the first-wave education reforms influence students' course-taking in high school? During the 1980s, national data show that course-taking in the core academic subjects did increase significantly, particularly in mathematics and science. Yet the extent to which this increase was driven by states' new graduation requirements remains unclear.

Also at issue is the impact of these reforms on academic engagement. Critics have suggested that minimum-competency exams and stricter coursework requirements will indicate to students that learning for its own sake is not worthwhile. In particular, these standards may encourage otherwise high-achieving students to avoid challenges and simply choose the path of least resistance to satisfy their requirements. Two decades ago, the authors of *A Nation at Risk* expressed a similar worry, suggesting that minimum-competency exams ran the risk of becoming "maximum" standards and thus lowering expectations for high-ability students.

To investigate the claim that course requirements and basic-skills tests may dampen the efforts of high-ability students, I analyzed whether higher graduation standards did indeed encourage students to take more courses in the core academic areas. I also investigated whether these requirements narrowed curricular effort by looking at students' course-taking in the visual and performing arts. To explore how these reforms may have influenced the intellectual engagement of students, I examined changes in time spent reading for pleasure, watching television, and doing homework.

Effects on Future Employment (Figure 2)

Minimum-competency tests and higher course requirements[a] for graduation raised the value of a high-school diploma and thereby increased students' probability of finding a job—even for those who dropped out of school. Black males benefited the most.



[a]High course requirements are defined as a high-school curriculum that includes at least three years of English, two of social studies, one of science, and one of math.
*Statistically significant at the 10 percent level
**Statistically significant at the 5 percent level
***Statistically significant at the 1 percent level

SOURCE: Author

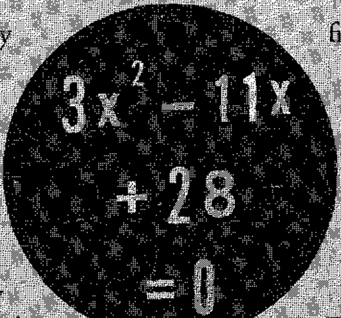
The data for these evaluations were created by pooling observations from two of the National Center for Education Statistics's major longitudinal studies: the sophomore cohort from the High School and Beyond study and the National Education Longitudinal Study of 1988. These surveys provide student-level data for nationally representative samples of 10th graders in 1980 and 1990—before and after the period when most first-wave reforms were implemented. Combining these data sets allowed me to adopt an analytic approach similar to the one I used to examine the effects of more stringent graduation requirements on educational attainment and employment.

My results uniformly suggest that the introduction of minimum-competency exams reduced course-taking in the four main academic areas, though these findings were statistically significant only in the sciences and in mathematics. This evidence suggests a lowering of intellectual aspirations among students who clearly exceeded the testing standards. However, an alternative interpretation is that schools and teachers redirected their efforts to respond to the needs of lower-performing students. Or it could simply reflect a reduction in curricular effort among those eventually encouraged by these reforms to drop out. But other evidence suggests that this is not the case. For example, the introduction of minimum-competency exams only increased the probability of dropping out among black males. But the reductions in course-taking associated with minimum-competency exams were more uniform across demographic traits and were particularly large for female students. Furthermore, the new tests also led to large and statistically significant reductions in credits for taking calculus, a margin relevant only for high-achieving students.

In contrast to minimum-competency testing, higher course requirements appeared to generate significant increases in the number of credits students earned in most of the core areas. For example, a high course requirement increased the number of credits earned in science by almost 0.4 Carnegie units, or 16 percent. Only in social studies was the estimated effect of a high course requirement statistically insignificant. These results suggest that higher standards contributed substantively to the upgrading of high-school curricula over this period, particularly in English and the sciences. The effect associated with higher course requirements, for instance, is equivalent to roughly 60 percent of the average growth in science credits over this period.

Were these gains in the core areas achieved at the expense of other valuable subjects, such as the visual and performing arts? The evidence is not entirely clear since the effects of first-wave reforms on students' participation in school arts and music activities were statistically indistinguishable from zero. However, I found that the first-wave reforms did adversely influence the amount of time students spent reading for pleasure, watching television, and doing homework. Higher course requirements were associated with large and statistically significant reductions in the amount of time spent reading for pleasure and doing homework and a corresponding increase in television watching. These results suggest that course-taking standards harmed the intellectual engagement of students. However, the apparent reduction in time spent on homework also raises the possibility that newly required courses were more demanding in prospect than in reality.

Higher standards contributed substantively to the upgrading of high-school curricula, particularly in English and the sciences.



Lessons

In the end, these findings indicate that the first wave of education reforms elicited many of the positive and negative consequences predicted by supporters and opponents. For example, both minimum-competency testing and increased coursework requirements led to reductions in educational attainment that were most pronounced among black students. Furthermore, minimum-competency testing led to some apparent reductions in the amount of courses completed, while increased coursework requirements appear to have had modest pejorative effects on the amounts of time students spend watching television, doing homework, and reading for pleasure. However, both types of reforms also increased students' subsequent likelihood of being employed. And higher coursework requirements were partially responsible for the apparent upgrading of high school curricula during this period. In light of these mixed results, what can the nation's experience with earlier reforms contribute to the current discussion of standards, testing, and accountability?

One lesson from our experiences with first-wave reforms is that standards-based reforms can generate a variety of costs and benefits, which can be compared only on subjective grounds. Some proponents of standards argue that we should implement targeted initiatives that soften the harsh consequences of standards by helping those at increased risk. However, any enthusiasm for that approach should be tempered by the knowledge that many of the state-level first-wave reforms were actually bundled with such efforts.

Our experiences with first-wave reforms also provide some suggestive evidence on the design of standards. In particular, these findings suggest that so-called "process" standards, such as higher course requirements, may be more effective than test-based standards, at least when testers seek only to establish a minimum level of performance necessary to graduate. While improved curricular standards yielded benefits in employment and in course-taking, minimum-competency testing had relatively few of the desired effects on educational attainment and early labor-market experiences. These results are consistent with the widely held perception that these test-based standards were often quite weak because of political pressures and the relatively easy and veiled manner in which they could be subsequently lowered. In contrast, newly introduced course requirements for graduation created binding standards for students and were also largely immune to political redesign.

Of course, whether these comparative first-wave results can be applied to the current generation of accountability systems is clearly open to conjecture. Perhaps the strategy of testing and tracking students' progress throughout their education will prove more effective than a single high-stakes exam at the end. But the early state-level experiences with minimum-competency testing provide a promising yet cautionary tale.

–Thomas S. Dee is an assistant professor of economics at Swarthmore College and a faculty research fellow at the National Bureau of Economic Research.

check the facts

by MICHAEL PODGURSKY

Fringe Benefits

There is more to compensation than a teacher's salary

Each year, the two national teacher unions, the American Federation of Teachers (AFT) and the National Education Association (NEA), release their surveys of public school teacher salaries across the nation. And each year, they take advantage of this opportunity to bemoan the condition of teacher pay. On the April 2002 release of the NEA's data, then–NEA president Bob Chase complained, "It's hard to convince someone to stay in the classroom when the salary is so low." Likewise, the AFT decried the fact that the "average teacher salary continues to fall well below the average wages of other white-collar occupations." The average teacher, according to the AFT, earned $43,250 during the 2000–01 school year, compared with an average of $52,664 for mid-level accountants; $71,155 for computer system analysts; $74,920 for engineers; and $82,712 for attorneys.

Of course, the AFT has chosen the comparison groups to make its best case. Where, one wonders, are the comparisons with journalists, registered nurses, assistant district attorneys, FBI agents, military officers, and other not-so-highly compensated professionals and public-sector employees? Shouldn't the average pay of a high-school English teacher be compared with that of writers and editors? One could make a case that the salaries of high-school physics or calculus teachers should bear some resemblance to those of computer system analysts, but does the AFT believe that the appropriate compensation benchmarks for 3rd-grade teachers are the salaries of engineers or attorneys?

Nevertheless, data from the NEA and AFT are highly influential. Indeed, the U.S. Department of Education collects few data of its own on this matter. For the most part it simply recycles these union data in publications like the *Digest of Education Statistics 2001*, a standard reference in which five of the six tables on teacher pay are based on union figures. On the whole, such data present a fairly accurate picture of teacher salaries at the national level and have some value for state-to-state comparisons. Yet they suffer from severe limitations when interest groups, policymakers, and pundits use them to make a point about how the nation values public school teachers.

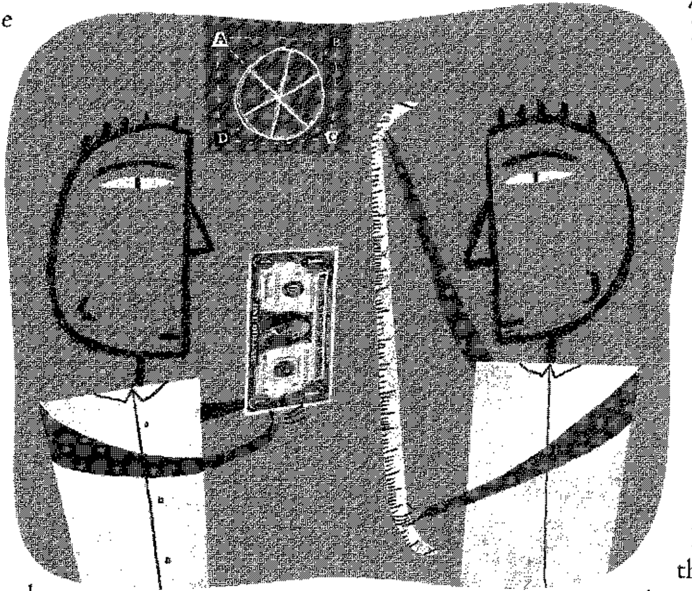


ILLUSTRATION BY JAMES YANG

Summers Off

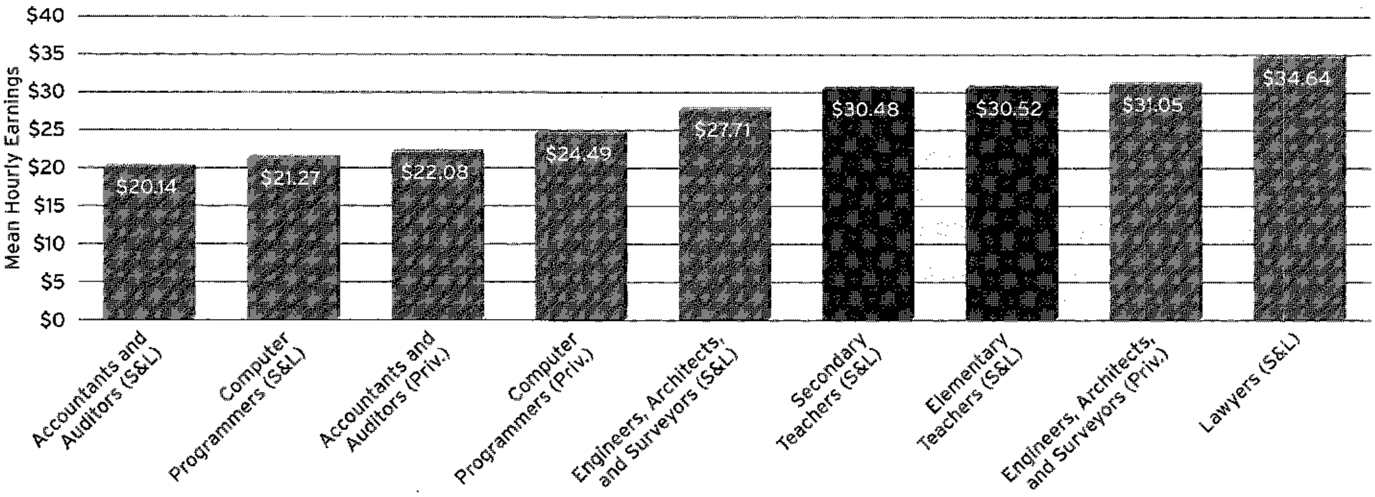
One facet of teaching that the NEA and AFT, in their data and in their public pronouncements, routinely fail to account for is the shorter workday and work year. In public schools, the median number of school days is 180 per year. Add half-a-dozen or so workdays for parent conferences, professional development, and planning, and the annual work year for most teachers is still shorter than 190 days. By comparison, an accountant or lawyer with two weeks of paid vacation and ten holidays or personal days will work 240 days annually—nearly 30 percent more days per year than public school teachers.

The typical teacher also has a shorter on-site workday than most other professionals. On average, teachers report being in school for fewer than 38 hours per week. This number rises to 40 hours if largely voluntary after-school activities such as coaching or club sponsorship are included. In fact, language limiting the number of hours that teachers are required to be in school is common in their collective-bargaining agreements, particularly in urban school districts. In the just-expired New York City teachers' contract, the contractual workday was just 6 hours and 20 minutes (including a 50-minute duty-free lunch). The new contract extends the workday by 20 minutes. In Chicago, the limit is 6 hours and 45 minutes, including a 45-minute duty-free lunch.

Of course, many teachers put in nights and weekends at home grading papers and planning for the next week. However, a job that permits relatively more work at home is typically more attractive (particularly to women with children) than one that requires a sim-

Solidly Middle Class (Figure 1)

Because of their shorter work year and workday, teachers actually earn more than many other educated professionals on an hourly basis.



*** S&L = employed by state and local governments; Priv. = employed by private firm**

SOURCE: Bureau of Labor Statistics

ilar amount of work time on site. And many other professionals bring their work home as well.

The combination of a shorter workday and work year means that the annual hours on the job for teachers are much shorter than in comparable professions. Consider Figure 1, which shows hourly rates of pay computed by the Bureau of Labor Statistics for a variety of occupations. By these calculations, only engineers, architects, and surveyors in private practice and attorneys earn more than teachers on an hourly basis.

The shorter workdays and work year make teaching an attractive occupation to those who wish to balance work and family needs. The shorter hours are especially helpful to women who want both a rewarding career and children, which helps to explain why roughly 75 percent of teachers are women and the share is increasing. An additional plum is that the on-site teaching hours match the schedule of a teacher's own school-age children. In short, when the kids are at home, so is mom (or dad).

Consider the "sick kid" challenge. In many professions it is very difficult to take unscheduled time off for a sick child or other family emergencies. In teaching, however, the "substitute teacher" solution is a routine part of school life. Indeed, in collective-bargaining negotiations, school administrators frequently complain of excessive absences among teachers. According to a recent U.S. Department of Education survey, during the 1999–2000 school year, 5.2 percent of teachers were absent on any given day on average. That translates into 9.4 days out of a 180-day school year. During the 2000–01 school year in New York City, the annual rate of absences reached 11.3 days per teacher. These rates are much higher than in other executive or professional employment. The Bureau of Labor Statistics reports that the absentee rate for managerial and professional employees is just 1.7 percent of annual hours.

Teaching is also family friendly in the sense that little or no out-of-town travel is required for successful job performance. Teachers may choose to attend out-of-town conferences, but such travel is unusual and not a condition of employment. Although I am not aware of any systematic data collected on this topic, out-of-town travel seems to be commonplace for many young professionals.

The expansion of opportunities for women during the past half-century is supposed to have lessened the attraction of teaching. And yes, the earnings of college-educated women in other fields have grown faster than the earnings of teachers in recent decades. However, the mix of nonteaching jobs that college-educated women hold has changed as well. In 1960, 58 percent of college-educated women not employed as teachers worked as secretaries or other clerical workers, and only 13 percent were "managers." By 1990 the clerical share had fallen to 30 percent while the share of managers increased to 35 percent (as did the shares of lawyers, accountants, and doctors). This shift from clerical to managerial and professional employment resulted in an increase in nonteaching earnings, but it also meant longer workdays and work weeks, greater responsibility (and stress), and, probably, less flexibility compared with teaching.

The bottom line is that teaching remains a job that makes it easier for parents to reconcile a career and family. Consequently we would expect female teachers to have more children—which is exactly what we see: among college-educated women aged 40 and younger, the average teacher had 2.1 children, versus 1.7 for other occupations. Of course, some of this difference may simply reflect the fact that teaching attracts women who like children and who would have been predisposed to have more children anyway. However, the fact remains that teaching is a profession that makes it less costly for these women to act on their preferences.

Starting Salaries

The most reliable data on pay and benefits for nonteachers are collected by the Bureau of Labor Statistics. Its economists and statisticians have a well-deserved reputation for maintaining very high standards for accurate and objective data collection on wages, benefits, labor force, and other economic data. Yet the AFT also presents tables and charts on earnings for new college graduates that rely on data collected by a private organization, the National Association of Colleges and Employers (NACE). Several times a year, NACE solicits reports on salary offers (not earnings) to graduating students. The AFT compares data from these NACE reports with scheduled starting salaries for teachers. In releasing the 2002 report, the AFT made much of the fact that "for new teachers, the $28,986 average beginning salary lagged far behind starting salary offers in other fields for new college graduates." By comparison, the AFT says, accounting graduates were offered an average starting salary of $40,779; sales/marketing majors, $40,033; computer-science majors, $49,749; and engineers, $50,033 (see Figure 2).

While these NACE data apparently have some value to placement officers on college campuses (who pay the hefty subscription fee), they almost certainly overestimate the earnings of nonteachers. First, these data represent salary offers to a very small sample of the total number of college graduates. The 2001–02 report, for example, was based on only 2,600 offers to students in fewer than 120 higher-education institutions. More important, this sample of offers is likely to result in a misleading estimate of the earnings of all college graduates, since only large businesses or employers are likely to send recruiting teams to campus. Most small colleges do not submit reports to NACE. At a large state flagship institution such as mine, the University of Missouri–Columbia, only the business and engineering schools report job offers to NACE. Moreover, for the Missouri business school, the number of graduates far exceeds the number of job offers reported by the placement office.

The AFT's data on starting pay for teachers must also be treated with some caution. Not all states collect such data, in which case the AFT estimates teachers' starting pay. In the most recent report, the AFT had to estimate the starting salary in 17 states. However, researchers have no way of judging the accuracy of the AFT's estimates, because they failed to document their methods.

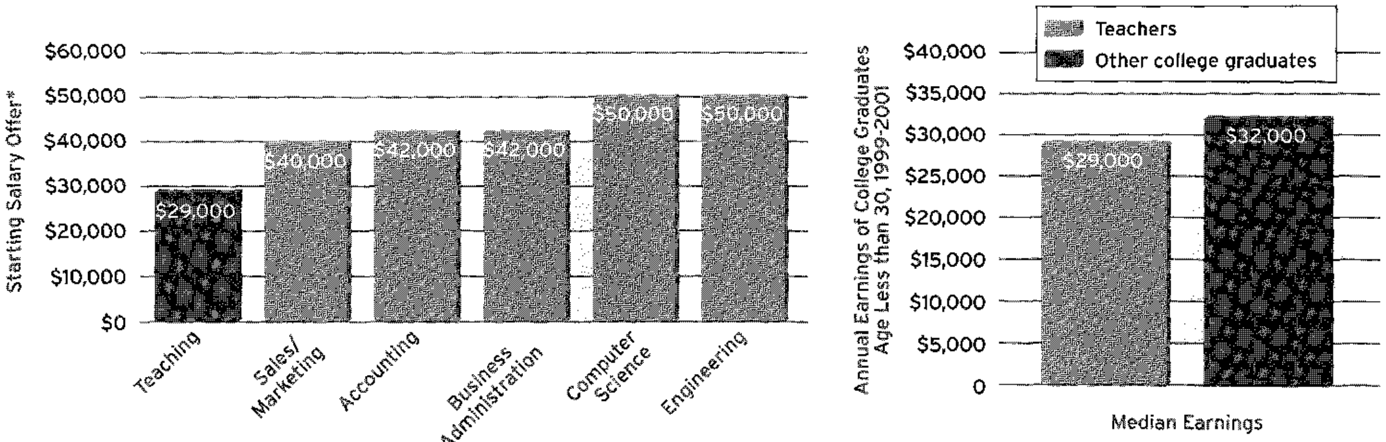
A much more accurate picture of the relative pay of recent college graduates can be gleaned from Bureau of Labor Statistics national survey data on annual earnings. I pooled three years (1999 through 2001) of March Current Population Survey data and computed full-time annual earnings for college graduates under 30 years of age whose highest degree is a bachelor's or master's degree. I made no adjustments for hours or weeks of work. Not surprisingly, nonteachers earn more than teachers (see Figure 2): on average, 32 percent more (nonteachers earned $36,996, versus $28,156 for teachers). However, the pool of nonteachers has many more very high earners than the pool of teachers, which pulls the average up. Thus a more representative indicator of the typical teacher and nonteacher is median earnings. In this case, the gap in earnings between nonteachers and teachers falls to just 10 percent ($32,000 versus $29,000). The latter gap is readily explained by the shorter workday and work year for teachers.

Fringe Benefits

Neither the AFT nor the NEA makes

Starting Salaries from American Federation of Teachers Survey (Figure 2)

The American Federation of Teachers uses data from the National Association of Colleges and Employers (NACE) to compare the starting salaries of teachers with those in other professions. However, the NACE data dramatically overstate the starting pay of college graduates. Data from the Bureau of Labor Statistics show a gap of only $3,000 in the median pay of teachers and other college graduates under the age of 30—a gap that is readily explained by the shorter workday and work year in teaching.



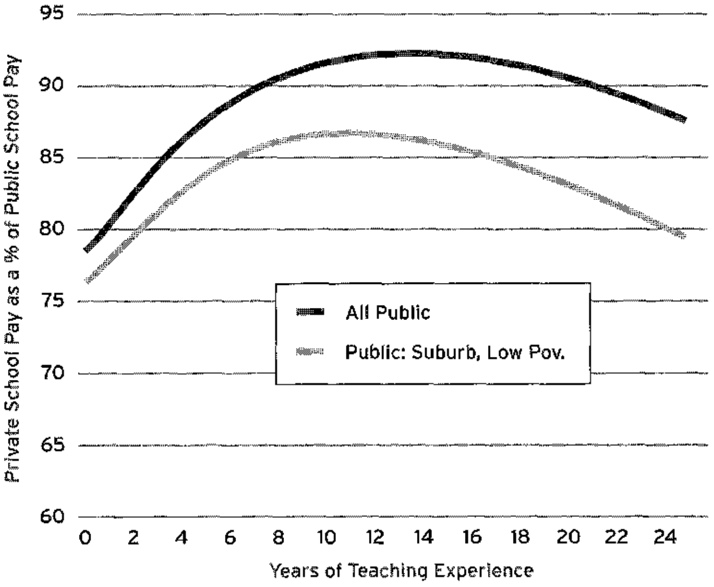
***Rounded to nearest thousand**

SOURCE: American Federation of Teachers, 2001, using data on starting pay of college graduates from National Association of Colleges and Employers (NACE)

SOURCE: March Current Population Survey, 1999-2001

Nonsectarian Private School Pay as a Percentage of Public School Pay (Figure 3)

Public school teachers fare much better than their colleagues in private schools. Salaries in nonsectarian private schools start at just 78 percent of those in public schools; climb as high as 92 percent of public school salaries in a teacher's 12th year; and decline thereafter. Private school salaries are even lower when compared with public school salaries in suburban, low-poverty settings.



SOURCE: U.S. Department of Education, 1999-2000 Schools and Staffing Surveys

any adjustments for the fringe benefits associated with teaching in a public school, thus masking an important part of total compensation. Unfortunately, published data do not permit precise comparisons of fringe (nonsalary) benefits for teachers and other professionals. For example, the Bureau of Labor Statistics reports fringe benefits for other professions, which include paid vacations. But nearly all teachers are on nine- or ten-month contracts and thus do not receive paid vacations. In addition, due to idiosyncrasies in the federal Social Security law, in some states teachers are included in state pension plans as well as Social Security, while in others they are not. Published Bureau of Labor Statistics data on benefits for "teachers" employed by state and local governments also combine full-time postsecondary teachers with K–12 teachers. This makes the comparison more difficult, but since public K–12 teachers account for more than 80 percent of the combined total, the data still provide some insight into the comparative benefits of teachers.

Employers' largest fringe benefit cost is retirement plans. Virtually all public school teachers are included in traditional defined-benefit plans in which teachers receive pension payments according to a defined schedule on retirement. These differ from defined-*contribution* plans, like the TIAA-CREF plans in most public and private colleges, in which the retirement benefits depend on investment earnings and saving rates and may vary from employee to employee. Public school teachers become eligible for pension benefits (or "vest") in these plans after five to seven years of employment. The contributions to these systems made by school districts or states are substantial. And because of the high turnover rates of teachers in their early years, these defined benefit plans in practice transfer wealth from young to more senior teachers.

The result is a system that permits teachers to retire earlier than they would if they were covered by Social Security or a conventional pension plan. For example, in the Missouri teacher pension system, a teacher who began teaching at age 22 and served continuously could retire at age 55 with 84 percent of her annual salary. In addition, her pension payments would be adjusted for inflation on an annual basis. Regular cost-of-living adjustments are unusual in private-sector defined-benefit programs. This teacher could also take employment in a new job and still collect her full pension benefits as long as the new employer was not a Missouri public school district.

National data show a similar pattern. In the 1994–95 Schools and Staffing Surveys, both male and female teachers who retired by the next school year averaged 59 years of age at retirement. By comparison, new retirees collecting Social Security retirement benefits have average retirement ages of 63.7 (men) and 63.6 (women).

The second largest fringe benefit cost (as a percentage of payroll) is health insurance. Health insurance coverage for public school teachers is nearly universal (more than 99 percent). The Bureau of Labor Statistics reports that health insurance benefits amount to 7.1 percent of hourly compensation costs for teachers (including postsecondary), but only 5.1 percent for professionals in private business. These same data suggest that the benefits provided to teachers are attractive relative to the private sector. For example, for individual policies, only 20 percent of health insurance plans for professional or managerial employees in medium and large private-sector firms are fully paid by the employer. The comparable share for teachers is 51 percent. Only 10 percent of medium and large private firms pay the full premium for family policies, compared with 29 percent in public school systems.

Private-Sector Teachers

Understandably enough, in comparing the salaries of public school teachers with those of other professionals, the AFT does not make what may be the most relevant comparison: between public and private school teachers. Perhaps this is why the levels of pay in private schools play such a small role in discussions of compensation in public schools. In areas other than K–12 education, personnel managers routinely use pay and benefits in the private sector as a benchmark in setting government rates. This holds for professional jobs such as lawyers, accountants, and nurses. Indeed, one important function of the compensation data collected by the Bureau of Labor Statistics is to provide private-sector as well as state and local benchmark data for federal wage setting. Public higher-education administrators are well aware of the level and structure of compensation for faculty in private institutions.

But public school districts rarely consult private school data in making their compensation decisions, even though 12 percent of teachers are employed in private schools and the two sectors compete for teachers. Mobility between the two sectors is extensive: 36 percent of full-time private and 13 percent of full-time public school teachers report some teaching experience in the other sector.

There are, however, some legitimate objections to public-private comparisons. First, many private schools have a religious orientation and are staffed by teachers of the same religious denomination. To the extent that such schools are advancing a religious mission, their teachers may be willing to work for less out of a religious commitment. Second, private schools are generally more selective in admissions than public schools, and, on average, their students are from households with higher socioeconomic status. To the extent that this results in better-behaved and more academically motivated students, it makes for a more attractive teaching environment in private schools.

Figure 3 compares public and private teacher salaries in a manner that attempts to address these concerns. First, I included earnings data only for teachers in nonsectarian private schools. In addition, I excluded private schools that have a special emphasis (special education, Montessori, Waldorf) and focused on schools that most closely resemble traditional public schools in mission. I also controlled for teachers' experience, gender, education level, region, and urban or rural status. As the graph shows, teacher pay in these private schools is consistently below that of public school teachers. Starting pay in private schools begins at 78 percent that of public schools, rises to 92 percent of public school pay by a teacher's 12th year, and declines thereafter.

Even with the above adjustments, a critic might argue that private school teaching is not comparable with public school teaching since the socioeconomic status of private school students is higher. In order to make the public schools more comparable with private, I excluded more than 90 percent of the sample of public school teachers and retained only public school teachers in suburban schools with little poverty (fewer than 5 percent eligible for free and reduced-price lunch). With these new criteria in place, the results look quite different. Interestingly, the shape of the pay function remains the same, but simply shifts down. Private school teachers now start at salaries that are 76 percent of their public school counterparts. This increases to 87 percent by their 12th year and declines thereafter. These results suggest that compared with the private sector, public schools overre-

Chicago Public Schools Salary Schedule, 2002-03 School Year (Table 1)

Schedule of annual salaries based upon a six-hour day during the school term of 40 weeks for regularly appointed members of the teaching staff.

LANE I: BACHELOR'S DEGREE

Step	Annual Salary	Pension Contribution	Total Compensation
1	$34,540	$2,420	$36,960
2	36,330	2,540	38,880
3	38,280	2,680	40,960
4	40,070	2,810	42,880
5	41,860	2,930	44,790
6	43,660	3,060	46,720
7	45,150	3,160	48,310
8	46,950	3,290	50,240
9	48,740	3,410	52,150
10	50,540	3,540	54,070
11	52,480	3,670	56,150
12	54,270	3,800	58,070
MAX	$55,890	$3,910	$59,800

LANE II: MASTER'S DEGREE

Step	Annual Salary	Pension Contribution	Total Compensation
1	$36,930	$2,590	$39,520
2	38,730	2,710	41,440
3	40,670	2,850	43,520
4	42,460	2,970	45,440
5	44,260	3,100	47,350
6	46,050	3,220	49,280
7	47,550	3,330	50,870
8	49,340	3,450	52,790
9	51,130	3,580	54,710
10	52,930	3,710	56,630
11	54,870	3,840	58,710
12	56,670	3,970	60,630
MAX	$58,350	$4,080	$62,430

SOURCE: Chicago Public Schools

ward high levels of experience. Why do they do this? In a recent study, Dale Ballou and I found that public school districts that are unionized and have a large share of high-seniority teachers are more likely to "backload" salary increases by adding additional steps to salary schedules or raising the rewards for seniority.

Benefits are lower in private schools as well. The median nonsectarian private school reports fringe costs (including Social Security) as 18 percent of payroll, while the comparable figure for public schools is 21.5 percent.

Salary Trends

Computing the change in mean teacher salaries from year to year can paint a misleading picture of the average teacher's situation. The reason is that public school salaries are set by schedules laid out in the teachers' contract. These pay schedules, which apply to all the teachers in a school district, base salaries on a teacher's years of experience and number of hours of graduate credits or graduate degrees. (See Table 1 for a typical salary schedule, from the Chicago Public Schools.)

When school districts raise the pay of teachers, they typically increase all the cells of these schedules by a fixed percentage, say 2 percent. Thus, if the education and experience of the average teacher in the district did not change from one year to the next, then average teacher pay would increase by 2 percent as well. However, that does not mean a typical teacher experienced a 2 percent pay increase between the two years. In fact, most teachers (except those who have "topped out" on the schedule) will receive a pay increase that is larger than 2 percent.

For example, a new teacher with a bachelor's degree hired in Chicago in the fall of 1999 earned $32,561. Between the fall of 1999 and the fall of 2002, the average starting pay for Chicago teachers rose a modest 6.1 percent. However, by the fall of 2002, that teacher hired in 1999 is now in her fourth year of teaching and her pay will have increased to $40, 071, or 23.1 percent in three years. If the 2003–04 salary schedule again rises by 2 percent, as it has over the past several years, then at the start of the 2003 school year her salary will have grown by 31.1 percent. If she earns a master's by that time, as many teachers do, her pay will have increased by 38.6 percent in four years—an average annual growth rate of 8.5 percent.

Given the steep tilt of salary schedules, changes in the experience mix of the workforce can be an important factor affecting changes in average teacher pay from one year to the next. This experience-composition problem also arises in analyzing national trends in average teacher pay, since many school districts are hiring more inexperienced teachers in response to rising enrollments, falling class sizes, and the retirement of older, more highly compensated teachers. According to U.S. Department of Education data, the median number of years of full-time teaching experience for public school teachers fell from 14 during the 1994–95 school year to 12 in 1999–00. Such a decline in the seniority mix of teachers tends to artificially mask the growth of average teacher pay and to bias comparisons of growth in teacher pay with pay in other professions, where the workforce is typically aging. It also makes simple cross-section comparisons of average pay between school districts or states problematic. We cannot be sure whether average pay is 10 percent higher in state A than in state B because salary schedules are higher in A or the teachers are older.

Trends in teachers' salaries will also understate trends in teachers' compensation if the costs of fringe benefits are increasing faster than salaries. Health insurance is an obvious example. Bureau of Labor Statistics data show that insurance costs (primarily health) represent 5 percent of hourly compensation costs for private-sector professional specialty occupations and managers, whereas for teachers they represent 7 percent of compensation costs.

Between 1980 and 1998 private-sector wage and salary costs grew by 104 percent, whereas costs for health insurance grew by 337 percent (even *after* widely publicized efforts by private-sector employers to restrain costs by introducing copayments and joining HMOs). Unfortunately there are no comparable data on the rise in insurance costs for public school districts. However, even assuming that public school districts made equally strenuous efforts at cost containment and held their health insurance cost increases to 337 percent, to the extent that health insurance is a larger share of payroll costs for public schools than for other private-sector workers, the relative compensation of teachers has increased. In short, in comparing the earnings of teachers and nonteachers, one cannot examine only trends in salary and assume that fringe benefits have been constant.

As states and districts attempt to meet the "highly qualified" teacher requirements of the No Child Left Behind Act, it will become increasingly important to gather more complete and accurate data on the relative pay and benefits of teachers. These data should also be disaggregated to reflect the fact that teacher labor markets are local or regional, and not national in scope. Attention should also be paid to differences in the opportunities that teachers face in the world outside teaching: a high-school physics teacher's potential earnings are not the same as an elementary-school teacher's. Policy discussions in this area have relied almost exclusively on the salary data collected by the teacher unions. A valuable step forward would be for the National Center for Education Statistics, in collaboration with the Bureau of Labor Statistics, to expand data collection on the relative pay and benefits of public school teachers so that states and school districts can have objective, arms-length data on this important issue.

–Michael Podgursky is chairman of the department of economics at the University of Missouri–Columbia.

The Concord Review • National Writing Board
Emerson Prizes • National History Club
Varsity Academics®

The Concord Review

Since 1987, The Concord Review, a quarterly journal, has published 54 issues with 594 high school history papers by students from 42 states and 33 other countries. Essays of around 5,000 words, with (Turabian) endnotes and bibliography, on any historical topic (ancient or modern, domestic or foreign), may be submitted, with a completed copy of the Form to Accompany Essays, and a check for $40 made out to The Concord Review, to: The Concord Review, 730 Boston Post Road, Suite 24, Sudbury, Massachusetts 01776. We publish about 8% of the essays we receive. William R. Fitzsimmons, Dean of Admissions at Harvard College, has said: "We have been very happy to have reprints of essays published in The Concord Review, submitted by a number of our applicants over the years, to add to the information we consider in making admission decisions...All of us here in the Admissions Office are big fans of The Concord Review."

National Writing Board

The National Writing Board, founded in 1998, has now given an independent, unbiased assessment of high school history papers from 24 states, and sent each author a three-page report, with scores and comments, which she/he has asked us to send to college admissions officers (at 58 colleges so far), or simply could use as feedback on one of her/his best history research papers. History research papers of two lengths—around 2,000 words, or around 5,000 words—with (Turabian) endnotes and bibliography, may be submitted, with a notarized Submission Form and a check for $65, made out to the National Writing Board, to: the National Writing Board, 730 Boston Post Road, Suite 24, Sudbury, Massachusetts 01776. Deadlines are November 1 and June 1 each year. The following colleges and universities now endorse this independent assessment service for academic writing: Amherst, Bowdoin, Claremont McKenna, Colgate, Connecticut College, Dartmouth, Duke, Eckerd, Emory, Georgetown, Hamilton, Harvard, Illinois Wesleyan, Lafayette, Middlebury, Northwestern, Notre Dame, Pitzer, Princeton, Richmond, Tufts, the University of Virginia, Washington and Lee, Williams, and Yale.

The Ralph Waldo Emerson Prizes

The Emerson Prizes ($3,000) have now been awarded to 31 students, published in The Concord Review, who have shown outstanding academic promise in history at the high school level since 1995. The ninth annual award ceremony was hosted by The Cliosophic Society, the History Club chapter at St. Mary's Episcopal School in Memphis, Tennessee on April 5, 2003. The four laureates this year are from Maryland, Massachusetts, Louisiana and Tennessee. Past laureates have come from Czechoslovakia, Colorado, Washington, D.C., Florida, California, Tennessee, Vermont, New Zealand, Utah, Massachusetts, Russia, Washington State, Connecticut, New Hampshire, Illinois, Japan, and New York.

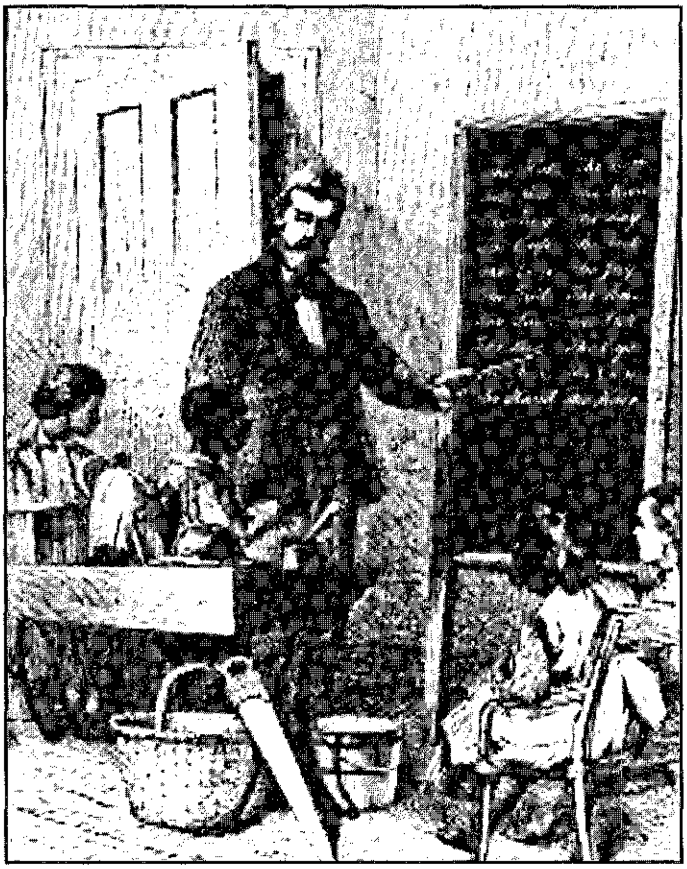
National History Club

The National History Club, founded in 2002, now has more than 800 members in high school history club chapters in eighteen states. Our Newsletter is online at http://www.tcr.org. Our goal is the promotion of the reading, writing, discussion and pleasures of history among high school teachers and students of history. There is no fee for joining. If your high school would like to apply for a charter, please send the names and addresses of all of your members and the faculty advisor to: Robert Nasson, Executive Director, National History Club, 730 Boston Post Road, Suite 24, Sudbury, MA 01776, email: nasson@tcr.org.

http://www.tcr.org

The two different submission forms for The Concord Review and the National Writing Board may each be found on this website, which has had more than 195,000 visitors, and since 1997 has been mirrored on a server in Singapore for the use of history teachers in Asia. This site also has 55 sample essays, including all of the Ralph Waldo Emerson Prize winners from the last nine years, and information on the National History Club, now with chapters in eighteen states. If you have any questions, please log on and leave a message for Will Fitzhugh at fitzhugh@tcr.org, or call (800) 331-5007.

The Concord Review, 730 Boston Post Road, Suite 24
Sudbury, Massachusetts 01776 USA
fitzhugh@tcr.org (800) 331-5007 http://www.tcr.org



penn

1-800-326-9180

state

www.psupress.org

press

JGE: THE JOURNAL OF GENERAL EDUCATION

Marilyn Amey, Editor
Elizabeth Jones, Sr. Associate Editor
David Guthrie and Elizabeth Larsen, Associate Editors

"As a member of the university council working on a revision of undergraduate general education requirements, I found the articles in JGE of particular relevance. The timely studies encouraged by the editorial board promise to create a forum for lively exchange on educational issues." —Julia Herschensohn, University of Washington

For faculty, administrators, and policymakers, JGE is the professional forum for discussing issues in general education today. JGE addresses the general education concerns of community colleges, four-year colleges, universities, and state systems. Along with perceptive essays on the role of general education today, JGE features articles on:

- Innovative methods in teaching and assessment
- Profiles of exemplary general education programs
- Case studies of successful curriculum development efforts
- Reviews of books and monographs related to general education

Annual Personal Subscriptions: $28.50; Outside USA $37.00
Annual Institutional Subscriptions: $49.50; Outside USA $57.50
Student Subscriptions: $20.00; Outside USA $28.00
Air Mail: Add $36.00 per volume
Terms: Prepayment is required (credit card, check, or money order)
Journals Subscription Office: 814-863-5992

by JENS LUDWIG

The Great Unknown

Does the black-white test-score gap narrow or widen through the school years? It depends on how you measure

Educational Achievement and Black-White Inequality

By Jonathan Jacobsen, Cara Olsen, Jennifer King Rice, Stephen Sweetland, and John Ralph

National Center for Education Statistics, July 2001.

Through the 1960s, African-Americans earned much less than whites—even when their cognitive abilities (as measured by test scores) were similar. By the end of the century, however, many believed that employment discrimination had attenuated to such a degree that the gap in labor-market outcomes could be explained almost entirely by differences in test scores. Consequently, reducing the well-known gap between the test scores of black and white students is now seen as an important way to reduce economic and other forms of inequality. In response, the U.S. Department of Education's National Center for Education Statistics commissioned Mathematica Policy Research, a contract research firm, to find out whether 1) employers are now playing fair and 2) schools are doing their part in narrowing the black-white test-score gap. Accepting the challenge, Jonathan Jacobsen and his colleagues discovered that the answer to the first of these questions was both more hopeful and much easier to provide.

Few dispute the dreadful state of affairs that existed in the 1960s. One study found that the black-white gap in scores on the Armed Forces Qualifying Test in 1964 could account for only a quarter of the difference in wages between black men and white men. Quite clearly, widespread racial discrimination helped create a labor market that yielded fewer rewards for blacks than for whites of similar ability. In addition, many jobs did not require a high level of cognitive skill.

Today, cognitive skills and educational credentials are more valuable for workers in part because of changes in production technologies that demand more highly skilled employees. Moreover, the civil-rights legislation of the 1960s and 1970s seems to have caused discrimination in employment to decline. As a result, the black-white gap in academic achievement now seems to account for a sizable share of the black-white gap in wages.

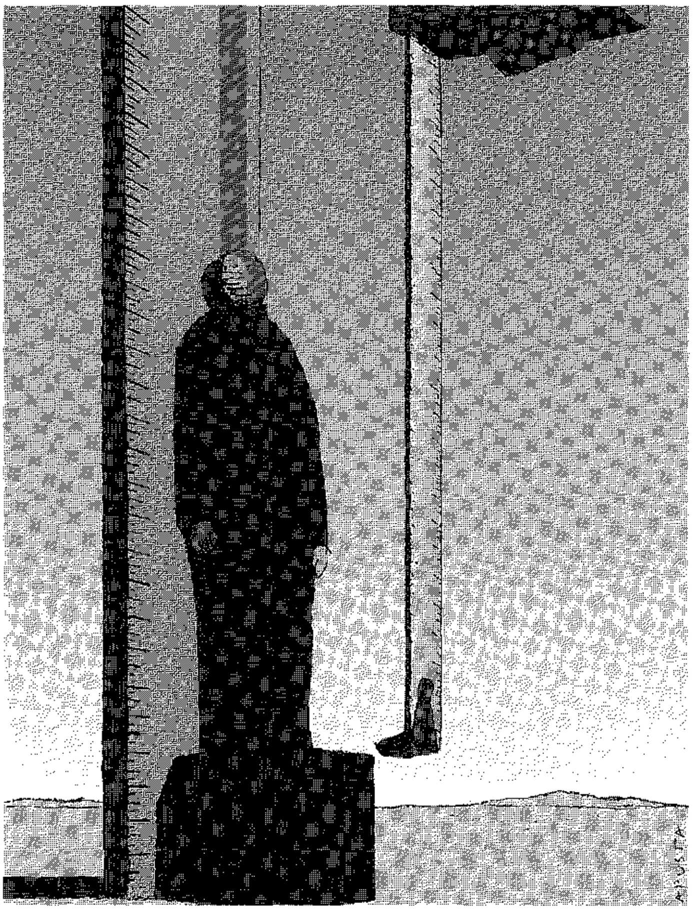


ILLUSTRATION BY JANUSZ KAPUSTA

Whether the gap in labor-market outcomes has virtually disappeared depends on how it is measured. In a widely cited 1996 study, economists Derek Neal and William Johnson showed that if test scores are not taken into account, white men's wages are 24 to 28 percent higher than those of black men. But these raw differences narrow substantially when differences in test scores are accounted for. Neal and Johnson, using data from the National Longitudinal Survey of Youth begun in 1979, showed that white men earned hourly wages that are only 7 to 10 percent more than black men with similar scores on the Armed Forces Qualifying Test. However, the gap in *annual* earnings between white men and black men with similar scores was about three times as large (roughly 30 percent). This is because white men are less likely to be unemployed, work longer hours on average, and are on the job more days of the year. Among women, wages and earnings were actually somewhat higher for blacks than for whites with similar test scores.

Confirmation

Using data from a variety of sources, including the National Longitudinal Survey of Youth, the High School and Beyond study, and the National Longitudinal Study of the High School Class of 1972, Jacobsen and his colleagues at Mathematica essentially confirm Neal and Johnson's findings, providing additional evidence that most of the remaining wage gap is due to differences in cognitive skills, as measured by test scores.

In the Jacobsen study, the hourly wages, employment rates, and annual earnings of black women were at least as high as those of white women with similar test scores and family backgrounds. The researchers also found that black and white men with similar scores and family backgrounds had similar wages and, contrary to the findings of Neal and Johnson, generally experienced similar employment rates and annual earnings as well.

The minor discrepancies between the findings of Jacobsen et al. and those reported by Neal and Johnson for male earnings and employment are presumably due to some combination of differences in sample definitions and outcome measures. For example, the Jacobsen study focused on respondents' earnings over a one-year period and excluded everyone who reported no earnings during that year. Neal and Johnson, by contrast, included in their sample all respondents who reported positive annual earnings for at least one year during the three years covered by their survey questions. Their strategy tends to include workers with more tenuous ties to the labor market than that used by Jacobsen et al. If blacks are likely to be less attached to the labor market than whites, it should come as no surprise that Neal and Johnson would estimate a larger black-white gap in earnings than that found by the Jacobsen report.

The black-white test-score gap now seems to account for a sizable share of the gap in wages.

In any case, the findings of Jacobsen and his colleagues lead to the same bottom line as previous research. Without a doubt, racial discrimination remains a problem to some degree in labor markets and other aspects of American society. Nevertheless, public policies that reduce differences in skills between blacks and whites before they enter the workforce could substantially reduce racial differences in labor-market outcomes.

Closing the Gaps?

Given this finding, the second question posed to the Jacobsen team is all the more interesting: Are schools gradually eliminating the test-score gap? Do black students and white students continue on in school at the same rate and become more similar in their test scores as they pass from one grade to the next?

One way to improve the skills of black students is to encourage them to stay in school for a longer period of time. In this respect, the Jacobsen report provides some heartening news. Jacobsen and his colleagues found that blacks are at least as likely as whites with similar test scores and family backgrounds to graduate from high school (or earn a GED) and attend college. Among college attendees, blacks are at least as likely to finish their degree program as whites with similar test scores and family characteristics. These findings confirm those from a variety of previous studies, which have shown that adjusting the data for students' socioeconomic status is usually enough to eliminate most or all of the difference between blacks and whites in educational attainment.

This arguably provides grounds for optimism for a number of reasons. First, ethnographic studies have suggested that black students are less inclined to try hard in school because their peers may view academic effort as characteristic of "acting white." Measures of persistence—how far students go in school—will of course miss the more subtle forms of academic effort. Nevertheless, the fact that blacks stay in school at least as long as whites from similar socioeconomic backgrounds seems inconsistent with the idea that the youth culture of blacks is much more anti-academic than that of whites.

Second, whatever problems colleges have in enrolling and educating students from disadvantaged backgrounds appear to be driven by issues of class more than by race. But what to do about the persistent effects of social class on educational attainment? If being a child in a low-income family mattered primarily because teens were forced to work or could not afford college, the government could reduce the black-white gap in educational attainment through more-generous tuition or income-assistance programs. However, research by economists Stephen Cameron and James Heckman suggests that the real problem may be that poor children simply don't do as well in school, a somewhat discouraging conclusion since sending families bigger checks to help with college is arguably easier than improving academic achievement in elementary and secondary schools.

Narrowing or Widening?

Staying in school is no longer enough, however. Students need to acquire cognitive skills along the way. How well are schools—and other forces outside of school—doing at reducing the black-white test-score gap as young people move from one grade to the next? At this point Jacobsen et al. are hard-pressed to find a convincing answer—and the one they do provide seems to depend on a number of analytical choices they made, some of which are debatable.

The best way to answer this question would be to follow a national sample of students through school from kindergarten through 12th grade. Surprisingly, the U.S. Department of Education has never done this. Instead, one cohort of students has been followed for a few years in elementary school, another cohort has been followed for a few years in middle school, and a third cohort has been followed in high school. Jacobsen and company are thus handicapped from the beginning in their search for an answer.

In an attempt to solve this problem, the Jacobsen team cobbled together data from the federal longitudinal "Prospects" study's Cohort 1 to measure children's test scores in 1st and 2nd grades (1992–93); Prospects Cohort 3 for scores in 3rd and 5th grades (1991–93); Prospects Cohort 7 for 7th and 9th grades (1991–93); and the National Education Longitudinal Study to measure 10th- and 12th-grade scores (1990–92). The available data enabled the researchers to derive measures of changes in the black-white gap for only a few years for each cohort's educational career. As a result, measuring changes in the gap over a longer period of time requires some comparison across cohorts.

The Jacobsen report, while hedging its bets, leaves the reader with the impression that the black-white test-score gap closes as students move through school. However, the evidence for the closing of the gap comes from problematic cross-cohort comparisons, not from the tracking of a specific cohort of students from one grade to the next (a "within-cohort" comparison).

Studying the more reliable within-cohort trends elicits less reason for optimism. In reading scores, two of the four within-cohort comparisons reveal a widening of the black-white gap as children move through school: between the 1st and 2nd grades, the gap grew one-third of a standard deviation larger, and another one-fifth of a standard deviation larger between grades 3 and 5. (There was no statistically significant change in the reading gap between grades 7 and 9 or grades 10 and 12.) In math, the within-cohort comparisons show the gap growing by two-fifths of a standard deviation between grades 7 and 9. (There was no statistically significant change in the math gap between grades 1 and 2, 3 and 5, or 10 and 12.)

Further complicating matters is the authors' choice of how to measure the black-white gap in each grade, which seems to drive much of what they find.

Another way to assess the performance of the education system with respect to racial disparities in achievement is to compare the subsequent performance of black and white students with similar initial scores. This approach is analogous to comparing the wages of blacks and whites with similar test scores, as discussed above. Again using the more reliable within-cohort comparisons, Jacobsen and his colleagues found that in both math and reading a black-white gap was virtually always present, even for students whose scores were similar just one or two years earlier.

Only in the across-cohort comparisons did the Mathematica researchers find much evidence to suggest that the black-white test-score gap may narrow during the school years. These comparisons seem to suggest that the gap in math scores narrows through elementary school, with some (but not all) of this gain lost during the middle-school years. In reading, the across-cohort comparisons suggest a narrowing of the gap through 7th grade, after which the gap remains stable.

The across- and within-cohort comparisons thus yield different stories about whether the reading gap narrows or grows as children progress through school. Which trend should be believed? An obvious problem with the across-cohort comparisons, Jacobsen and his coauthors note, is that each cohort of students may exhibit different characteristics, making cross-cohort comparisons difficult. In other words, this year's 3rd graders are not so easily compared with next year's 3rd graders; in one, cartoons like Pokémon may have been all the rage, while reading Harry Potter books might have caught on the next year. Unfortunately, the degree to which the report's findings can be explained by these kinds of across-cohort differences is not easy to assess. Data from the National Assessment of Educational Progress suggest a fairly complicated pattern of changes in the black-white gap across cohorts that depends on the subject area and the age at which the tests are administered.

Further complicating matters is the authors' choice of how to measure the black-white gap in each grade, which seems to drive much of what they find. To wit, the black-white gap in each grade is expressed as a fraction of the standard deviation in test scores observed in the 8th grade. Setting aside the question of whether 8th-grade tests can be compared with the lower-grade tests in the Prospects study and the higher-grade tests in the National Education Longitudinal Survey, a more basic problem with this measure has to do with the statistical properties of test scores.

Say that a large population of black and white 5th graders is given a standardized test. Each group's scores will form a distribution with a certain average and a standard deviation—a measure of the scores' range. Since white students score higher than black students on average, let's say that the average white score is 100, while the average score for black students is 80. Each distribution has a standard deviation of 20 points.

Now say the same group of students is tested again in 8th grade, where the average score of black students turns out to be 90, versus an average of 100 among white students. The difference in average scores is now 10 instead of 20. The test-score gap appears to have narrowed. However, imagine that the standard deviation of these distributions has also been reduced to 10 points, meaning that there is less variability than in the scores of 5th graders.

Has the black-white gap narrowed from 5th grade to 8th grade? The approach adopted by Jacobsen would suggest that it has: measured as a share of the 8th-grade standard deviation, the black-white gap has declined from two standard deviations to one. But this is not the most informative approach to measuring the gap. Notice in my example that the average score of black students lies at the same point in the white distribution in both the 5th and the 8th grades: around 75 percent of white students score higher than the average black student in both grades. If we care about how much the two distributions overlap each other, we should calculate what is known as an "age-standardized" measure of the black-white gap, defined as the difference in the black and white average scores in a given grade divided by the standard deviation in scores *in that grade*. Applying this approach to my example, one would conclude that the black-white gap has not changed across grades.

When University of California at Los Angeles professor Meredith Phillips and her colleagues compared age-standardized measures of the black-white gap across grades, they found that the math gap widened by around one-fifth of a standard deviation between the 1st and the 12th grades, while the reading gap did not change much. Unfortunately, given the limitations of the available data, Phillips also relied on across-cohort comparisons that are vulnerable to the problems mentioned above. Given this currently unavoidable problem, I remain agnostic on whether the black-white gap changes very much one way or the other as children progress through school.

In reading, two of the four more reliable comparisons reveal a widening of the black-white test-score gap as children move through school.

What Would Be Learned?

Suppose that enlightened policymakers eventually fund the type of longitudinal study that would enable the tracking of changes in the black-white test-score gap from 1st grade to 12th grade for a single cohort of students—precisely the type of study Jacobsen and his colleagues call for. It is still not clear what would be gained from such an exercise.

One motivation for studying the change in the black-white gap across grades has been to test the hypothesis that the difference is primarily genetic in origin, since evidence that the black-white gap changes in magnitude over time would suggest a role for environmental influences. But evidence for substantial environmental effects on test scores has been available for some time, including evidence of large IQ gains over time within a number of countries (the so-called Flynn effect, named for the psychologist James Flynn). Moreover, genetic influences need not have the same effect on test scores throughout the lifecycle, since any genetic effects may be amplified over time if they also affect the environmental influences that people seek out for themselves. While there are several compelling points against the genetic argument, variation in the size of the black-white test-score gap across grades is not high on the list.

Another motivation for tracing these changes is to see whether the nation's education system is contributing to equal opportunity. If no reduction in the gap can be detected, one has reason to doubt that schools are having their desired impact. Of course, it may be argued that schools are counteracting socioeconomic factors that would otherwise lead to ever-growing inequality (if, for example, family background becomes increasingly important for scholastic success as children age and homework becomes more important). These data do not help to answer that question. Unfortunately, without information about a single cohort of students throughout their years in school, it is not certain if the black-white test-score gap increases, decreases, or remains roughly the same through the school years. Underlining this fact may be the greatest contribution of the Jacobsen report.

Evidence about how the black-white gap changes as children move through school may help to identify which life stages deserve unusually intense study by policy analysts. However, the specific policy implications of changes in the black-white gap as children age are far from obvious. What no one disputes is that the average black child enters and exits the school system with lower scores than the average white child. This fact alone would seem to leave little doubt that more needs to be done to address this problem.

–Jens Ludwig is an associate professor of public policy at Georgetown University.

Marginal Impact

Time's too short for one kid at a time

Upstart Startup: Creating and Sustaining a Public Charter School

By James Nehring

Teachers College Press, 2002

Standards of Mind and Heart: Creating the Good High School

By Peggy Silva and Robert A. Mackin

Teachers College Press, 2002

Central Park East and Its Graduates: "Learning by Heart"

By David Bensman

Teachers College Press, 2000

One Kid at a Time: Big Lessons from a Small School

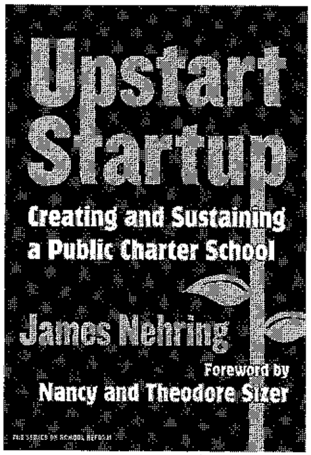
By Eliot Levine

Teachers College Press, 2001

Reviewed by Robin J. Lake

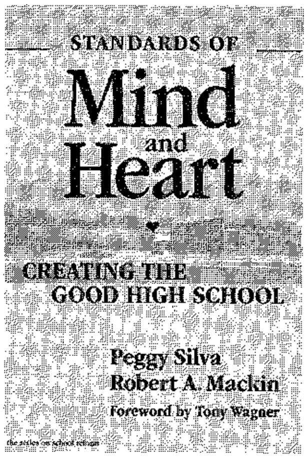
"Inspiration," "heart," "democracy," "models for national education reform"... so advertise the covers of four books from the series on small schools by Teachers College Press. True to promise, the authors deliver compelling stories of students once on the brink of academic failure who have thrived in four small schools, all based on principles from the Coalition of Essential Schools. Tony with Down's syndrome, Tamika from a broken home, and others similarly afflicted became engaged in learning thanks to superhuman efforts from teachers and student-centered learning.

These stories illustrate the argument in favor of creating smaller schools: students get personal attention, individualized learning plans, and grades based


Upstart Startup
Creating and Sustaining a Public Charter School
James Nehring
Foreword by
Nancy and Theodore Sizer
THE SERIES ON SCHOOL REFORM

on long-term projects and performance-based graduation requirements. All of these schools, for example, require "senior exhibitions" for graduation, where students present project results to a panel of outside experts. Most of these schools also regularly collect portfolios of student work in an attempt to go beyond standardized test results and provide richer measures of achievement.

This all makes small schools sound very attractive. But if small schools are to live up to their promise as national reform models and broader remedies for inequity, inspiration and heart may not be enough.



As these books demonstrate, operators of small schools share good ideas about how to make their own schools work, but they regularly ignore the fact that they have escaped from a system that prevents others from doing the same thing. Lots of educators who decide to create small schools, often by negotiating special arrangements with the local school district, do so because they don't have any confidence that the system can be changed in any real way. Their solution has been to try to save as many kids as they can get their hands on. But they act as if anyone could do what they have done. Their denial of their own special circumstances lets them ignore—and even oppose—the system-wide changes that must occur if their valuable ideas are to be widely replicated. They rail against state standards and accountability systems, but offer few practical alternatives for ensuring that *all* public schools perform at a high level. Any serious reform effort must articulate a credible strategy for replicating successful schools.

Reinventing the Wheel

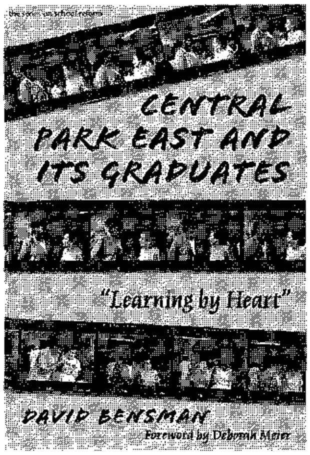
The strength of these four books is the honest lessons they offer people who are interested in starting new schools. *Upstart Startup* is the story of a Massachusetts charter school founded by Ted and Nancy Sizer (he the founder of the Coalition of Essential Schools). Hand-picked veteran teachers staff the school, and a constant stream of Harvard graduate students scrutinize its workings and volunteer their time. If any school should succeed, this is the one. Yet author James Nehring, the school's first principal, tells a story of utter chaos in the school's first year, including scheduling meltdowns, administrative blunders, and a failure to link course objectives with established graduation outcomes.

In *Central Park East and Its Graduates,* David Bensman describes similar early mistakes and regrettable incidents in the famous small school Central Park East. Now more than a decade old, at various times the school tried consensus-based governance and failed, struggled to get teachers to collaborate on instructional strategies, and butted heads over race and student favoritism.

The majority of new "small" schools are likely to make similar mistakes. The lure of building from scratch and deliberating until every teacher is happy with every decision is strong for small-school founders. And they all seem doomed to learn the same lessons in isolation. School founders have few places to turn for advice about how to avoid disaster. School districts simply do not have the expertise to help—or, worse, actively get in the way—and potential sources of technical assistance like the Coalition of Essential Schools are thinly funded and reluctant to intrude.

Inexplicably, even when lessons, resources, and help are available, school founders resist the notion of replication. Nehring's attitude is common. He argues that efforts at replicating his program are ill advised—not even a school's scheduling policies should be imitated. Those interested in starting their own schools are told not only to expect but also to embrace the "muddle through" strategy. Nehring argues that school reform is necessarily this messy and irreducible. By contrast, the entrepreneurial founders of the Met School in Rhode Island (as described in *One Kid at a Time*) have started a nonprofit to seed other Met-like schools around the country and are developing materials and assistance to make that happen.

Resistance to the idea of replication is a romantic notion that the small-schools movement and disadvantaged students cannot afford. If every school starts from scratch, the prospects are grim for reaching more than a scattered handful of students in every city. With foundations and governments willing to invest in the scaling up of promising schools, it is all the more disappointing when the books in this series fail to help answer the question of how to get beyond the boutique school and "mass produce" many smaller, personalized programs without losing the essential elements that make these schools a success?


the series on school reform
CENTRAL PARK EAST AND ITS GRADUATES
"Learning by Heart"
DAVID BENSMAN
Foreword by Deborah Meier

Peggy Silva and Robert A. Mackin do believe that their small school in New Hampshire can be replicated, but their faith is misapplied. The story of Souhegan High School in New Hampshire surely describes what high schools should strive to become: oriented around creativity and questioning, designed to ensure that students are known as individuals and will be aggressively targeted for intervention if they begin to fall behind. The implication is that every large urban high school can and should be like Souhegan. Largely ignored is the fact that Souhegan was designed as a new school, not a conversion from an existing program, and in the process attracted some of the best teachers in the country.

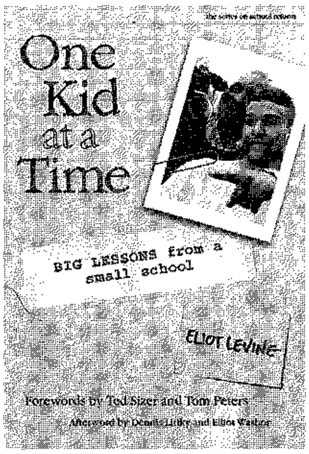
In fact, all of the schools in this series were started as new schools, chose their teachers and are chosen by families, and have unusual operating autonomy. It is probably no coincidence that two of the four schools are charter schools. Most advocates of small schools acknowledge privately that operating as a new school and allowing students to choose to enroll are essential elements of their success. However, these realizations are usually buried in the texts of books about successful schools. Policymakers and educators are allowed to imagine that good schools can be woven out of existing bad schools and that the difficult politics of transforming schools and creating choice for families and teachers can be avoided.

There are certainly lessons about scale-up hidden in these books, and they should be used. School incubators like those running in Dayton, Sacramento, Oakland, and other cities, and nonprofit providers of technical assistance could analyze small-school start-up lessons and train new school founders to avoid common mistakes. Districts could hire staff with expertise in supporting small-school development. More schools could imitate the Met by documenting their policies and procedures to make it easy for other schools to copy what they've done well. Those who fear that replication efforts will produce cookie-cutter schools are right to worry, but they also have an obligation to encourage the creation of franchises that account for local context and preferences but still maintain quality controls. The need is too great to build one school at a time.

Content Is ... Queen

A fundamental—and particularly controversial—aspect of the small schools described in these books is their resistance to prescribed curricular content. According to the authors, students in these schools came away with process skills adequate to allow them to become "lifelong learners." What about learning major historical events, the U.S. Constitution and Bill of Rights, basic knowledge of life sciences, and foreign languages? Students may or may not know about these things without a defined curriculum. It all depends on whether the topics fit their interests at the time. Graduates of these schools supposedly come away with the ability to seek out whatever content is needed to address academic questions as they arise. The theory is that kids don't retain knowledge unless the subject interests them, so why bother trying to get them to remember events and dates that they could easily look up in an encyclopedia if they ever have the need?

The authors make no bones about their disdain for those in the "content" camp, who wonder why a student should be able to graduate without knowing anything particular about history or math. Nehring, in *Upstart Startup*, tackles this debate most directly in his chapter titled, "Is Shakespeare Essential?" Nehring's point is this: of course his students are exposed to content through their project-based learning, but it would be futile for people at the school, district, or state level to try to agree on what content should be required. Nehring derides, for instance, the Massachusetts charter school inspection team that dared to single out the fact that his students didn't seem to know basic historical facts.


the series on school reform
One Kid at a Time
BIG LESSONS from a small school
ELIOT LEVINE
Forewords by Ted Sizer and Tom Peters
Afterword by Dennis Littky and Elliot Washor

Souhegan High School is an exception; graduates are required to complete a core curriculum in basic areas (math, science, literature, and history). The Met School, by contrast, exhibits the extreme, with no classes, just project-based learning. The other schools teach some content, but it is integrated in teacher-crafted, theme-based modules. Despite their rhetoric, most small-school staffs clearly value content to some degree. Unfortunately, these books do not offer much new or constructive insight on how schools that emphasize process skills can also be sure their students come away with a sufficiently strong content base. In response to outside criticism, Nehring begrudgingly proposes new Essential Content to complement the school's Essential Skills. The content list is broad enough to accommodate teachers' judgment and creativity, and is a good and potentially useful model for basic content standards in other Coalition schools. What a shame that innovations like this are so rare and are treated as embarrassing footnotes in these books.

A more middle-ground perspective from small-school founders on the issue of content would be refreshing and productive. Surely we can agree, for instance, that skill in finding and critiquing material is valuable, especially in today's information-overloaded Internet world. Surely we can also agree that one's ability to interpret and critique those materials is limited without knowledge of facts.

If small schools are to become broader remedies for inequity, inspiration and heart may not be enough.

Standard Hostility

State-mandated standards and tests are not popular in schools where every teacher individualizes content for every student. Not surprisingly, these authors devote considerable space to standards bashing. The arguments range from the persuasive (the standards are too prescriptive, are too broad, or divert attention from teaching of higher-order skills) to the absurd (the tests are beneath us). Dennis Litky (cofounder of the Met School) argues:

> *I know that states need to hold schools accountable in some way. . . . I'm not asking to be excused from that. What I disagree with is having one way to evaluate every student. Every state says they want students to become good thinkers, problem solvers, and citizens, but then the tests they give have almost nothing to do with those things!*

There may be something to this argument, but the solution Litky presents is absurdly subjective: every school should be able to devise its own standards and ways of showing it is achieving what the state claims to value. He writes, "We could videotape every exhibition and the state could randomly select and assess a sample of them." Litky and his fellow small schoolers can make these claims more reasonably than other schools, since most small schools are schools of choice. Thus they face an even tougher form of accountability than state standards and tests: the ability of parents and students to leave the school if they aren't satisfied.

Nevertheless, whether or not the authors share legitimate complaints about the constraints that standards place on their professional judgment, the "leave us to our rubrics" attitude marginalizes the small-schools movement. The authors offer thin evidence of whether their schools are in fact teaching and assessing content knowledge and skills in the ways they promise. All four books offer reasonably strong evidence that these schools are keeping kids from dropping out, getting them into college, and helping them do well on standardized reading tests. But descriptions of school-based outcomes and assessments in these schools raise troubling doubts. Litky again:

> *Do we require a minimum proficiency level? I'd like to say that we wouldn't graduate a student with a 6th-grade reading level. But on the other hand, there aren't even any measures that assess reading ability very well.*

Even if we buy the idea that teachers at the Met are so talented that they should be trusted to demand high performance from every student, what about schools with less skilled teachers and lower expectations? These are the schools that standards, standardized assessments, and state accountability systems are designed to identify and fix. By failing to take the need for states to address this problem seriously, small-schools advocates seem to ignore the broader issues of social justice that they work so hard to address in their own schools.

Instead of devoting so much energy to dismissing the standards movement, small-schools founders and advocates would do well to engage the discussion and help refine or redefine state standards and statewide accountability systems in the name of equity. Schools based on inspiration and heart may be enough for reformers who are willing to settle for localized quality and impact, but the small-schools advocates I know aspire to a lot more. They believe true equity depends on universal access to good public schools. For that to happen, small schools must mature into a stronger movement by formulating serious proposals for assessments and accountability, throwing their vocal political support behind school choice, and insisting on new, innovative efforts to reduce the learning curve for those who wish to start small schools and their overseers.

–Robin J. Lake is associate director of the Center on Reinventing Public Education at the University of Washington.

Study Abroad

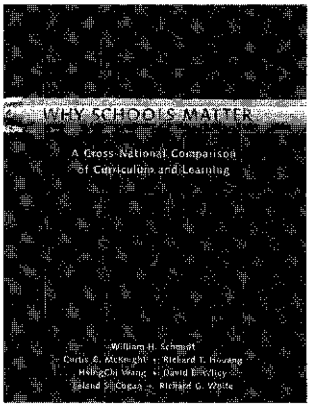
What TIMSS Teaches Us

Why Schools Matter: A Cross-National Comparison of Curriculum and Learning

By William H. Schmidt et al.

Jossey-Bass, 2001, $29; 400 pages

Reviewed by Stephen P. Heyneman



The International Association for the Evaluation of Educational Achievement (IEA) began with a chance meeting between Torsten Husen of the University of Stockholm and a group of University of Chicago researchers in the mid-1950s. The scholars at Chicago's Comparative Education Center, C. Arnold Anderson, Mary Jean Bowman, and Ben Bloom, believed that the whole world should be seen as an educational "laboratory." The first results from IEA's international surveys appeared in 1964. Since then, there have been 29 IEA-sponsored cross-national studies. The most famous of these was the 1995 Third International Mathematics and Science Study (TIMSS), the main subject of *Why Schools Matter.*

U.S. schools are characterized by too much breadth and not enough depth–in more familiar words, curricula are a "mile wide and an inch deep."

Up to 45 nations participated in at least one aspect of this cooperative study of schools, teachers, students, and learning in math and science at three age levels, making it perhaps the largest survey ever undertaken in the social sciences. As in previous surveys, the results suggested that the typical U.S. student knows less math and science than do students in many other industrialized nations. What TIMSS has taught us is that these differences are compounded over time: children in the United States drop lower in the rankings as they age.

One explanation is that curricula in U.S. schools are characterized by too much breadth and not enough depth—in more familiar words, they are "a mile wide and an inch deep." William Schmidt and his colleagues, the authors of *Why Schools Matter*, are largely responsible for this notion. They found that school systems in the United States cover a broader range of curricular topics than any other country save Switzerland. However, Swiss textbooks are more selective than American texts, making the U.S. curriculum the most "splintered" in the world.

Unlike the national curricula used in Europe and East Asia, curricula in the United States are uncoordinated, with insufficient attention paid to what students learn as they progress from grade to grade. In a highly mobile nation like the United States, this is particularly nettlesome when children change schools and discover that their new school district teaches geometry in 7th grade, while they took it the previous year.

Curricular variation complicates international comparisons as well. The TIMSS test for 13-year-olds included 44 different mathematic concepts. Some of these concepts weren't taught in all nations, while many others were taught at one age level but not at another or in different sequences. Does this mean that the tests were unfair? No. In fact, even when there were differences in coverage and sequence, these differences turned out to yield important information. It may be the case that children in the Netherlands are asked to move at a quicker pace, with a more logical sequence of topics, and thus learn more mathematics by the time they graduate. This information has proven useful for U.S. policymakers as they search for ways to upgrade the education system.

Students can even vary from nation to nation in their approach to these tests, further complicating comparisons. The international tests are low stakes in the sense that no consequences are attached to a student's performance. However, in some cultures, students may view all testing as a challenge to be taken seriously. In other cultures, students may not care in the slightest how they perform unless it "counts."

The Size of Minnesota

What *Why Schools Matter* fails to say is that countries themselves are often an arbitrary choice as a unit of analysis. For instance, the United States has a range of demographic, cultural, and social-policy differences that make a comparison with Sweden, a country with a population the same size as Minnesota, somewhat artificial. If other factors were used in defining a legitimate unit of analysis—size, wealth, and school functions—states might be the only units meeting the criteria in the United States.

Why Schools Matter also fails to discuss—mostly because it was not included in the TIMSS surveys—how teachers' knowledge and skills differed across countries and what effect that might have on achievement. Nor did the TIMSS study have any data on the amount of money nations invested in education. That TIMSS paid no attention to spending was no accident. Decisions concerning which data would be collected were made by IEA committees in which experts had to reach consensus. The committees themselves were overloaded with specialists on subject matter and methodology. No education economist, and few education sociologists, sat on a TIMSS committee. When the time came to decide on the focus of TIMSS, the subject-matter specialists had a virtual monopoly. Consequently, the measurements of the social contexts in which learning takes place are superficial. Schmidt and his colleagues do take a brief and superficial detour to see if their results change when they adjust the data to account for a nation's per-capita income (they did not change).

The lower a nation's GDP, it appears, the more influence the school, as opposed to socioeconomic status, seems to have. This has been a principal rationale for the investment in school quality by international development agencies.

Nonschool factors such as a child's social background seem to make a difference on the TIMSS tests at all levels. The book does not ask, however, whether this difference is the same in all countries or whether it varies in some systematic pattern. Research performed by William Loxley and myself in 1983 found that the influence of a student's socioeconomic status on achievement varies from one nation to another. The lower a nation's gross domestic product (GDP), it appears, the more influence the school seems to have. This finding has been a principal rationale for the investment in school quality by the World Bank, USAID, UNESCO, and many other international development agencies. Although recent re-analyses have challenged the strength of the earlier findings, the fact remains that the influence of socioeconomic status on achievement is by no means uniform across nations, age/grade levels, gender, and subject matter.

Why Schools Matter may be of marginal interest to general readers, but it is an excellent book based on a study of unprecedented sophistication and quality. It tells us that if U.S. schools are to improve, much more attention must be paid to the depth and focus of their curricula.

–Stephen P. Heyneman is a professor of international education policy at Vanderbilt University.

education matters to me

by KRISTIN KEARNS JORDAN

Meet Mr. Shannon

At Bronx Prep, a master teacher shares his expertise

At the Bronx Preparatory Charter School we enjoy the talents of a master teacher, Frederick Shannon. He had been teaching 5th grade in the classroom next door to our principal, Marina Bernard Damiba, when she began her career through Teach for America ten years ago. As a new teacher, Marina had camped out in his classroom during free periods and absorbed every nuance of his technique. Then he retired. But as Marina moved into leadership positions, first at the KIPP Academy and now at Bronx Prep, she coaxed him out of retirement.

Shannon and teacher Tara Kelley.

"You need to coach the next generation," she implored.

During his 25 years in teaching, Mr. Shannon developed the "Shannon Plan" of effective instruction. It works. He knows how to teach children and how to coach fellow teachers. He assists our young staff members with their lesson plans and visits their classrooms to offer suggestions for improvement. He demonstrates model lessons in math and reading. Our teachers are becoming more and more effective, thanks to Mr. Shannon's commitment and expertise and to the tailoring of their training to the context of their own classrooms.

Some teachers contemplate the aggravation of becoming certified and bolt to private schools.

To the new teacher who is struggling to get students settled at the beginning of class, he suggests the time-tested "do now," a silent opening exercise in which the teacher writes an assignment on the board before students enter the classroom. Students see the assignment under the "do now" heading and immediately focus on the task at hand.

Other techniques reflect his unique style of making lesson components more visually or physically compelling. With his support, our 6th-grade reading teacher has developed her signature hamburger lesson. Dressed as a hamburger, one of her students models topic and closing sentences that give the paragraph its structure (the bun) as well as the meaty support sentences and details inside. In another lesson, to help students remember that plot = conflict + resolution, teachers take students through a series of Shannon's hand signals that are part boxing, part mime.

Unfortunately, the state of New York does not recognize the holistic, integrated, practical training Mr. Shannon provides as valid preparation for a state teaching license. Currently, schools like ours, in an area of documented teacher shortage, are permitted to hire teachers with one-year emergency licenses. But this allowance will be phased out by 2004. Our teachers will then have to be certified. So each evening they scramble off to schools of education and to on-line learning programs to get the credits they need to stay in the business of teaching.

Credits are expensive. They also take time away from planning lessons, and they usually have little to do with the day-to-day reality of educating kids. Course offerings like the "Environments of School" may be intellectually stimulating, but what new teachers really need is practical, hands-on experience. Some teachers contemplate the aggravation and expense of becoming certified and bolt to private schools. Countless others evaluate the field's barriers to entry and pick another profession.

Good principals, like managers everywhere, know how to evaluate potential teachers. First, recruit well-educated candidates. Then, give them a class of 25 students for 45 minutes and see what the students learn. Call references to assess their ability to work with a team. Get beyond first impressions: meet with them several times before making the offer.

In an environment of choice and accountability, a principal's commitment to students should provide all the motivation necessary to select the best candidates. Teaching does not pay well, but there is nonetheless a pool of talented, motivated people eager to educate the next generation. Why is the state working so hard to narrow that pool?

–Kristin Kearns Jordan is the founder and executive director of the Bronx Preparatory Charter School.

PHOTOGRAPH COURTESY OF KRISTIN KEARNS JORDAN

an assessment by the koret task force on k–12 education

Our Schools & Our Future

...are we still at risk?

an assessment by the koret task force on k–12 education
Our Schools & Our Future
...are we still at risk?
edited by
Paul E. Peterson

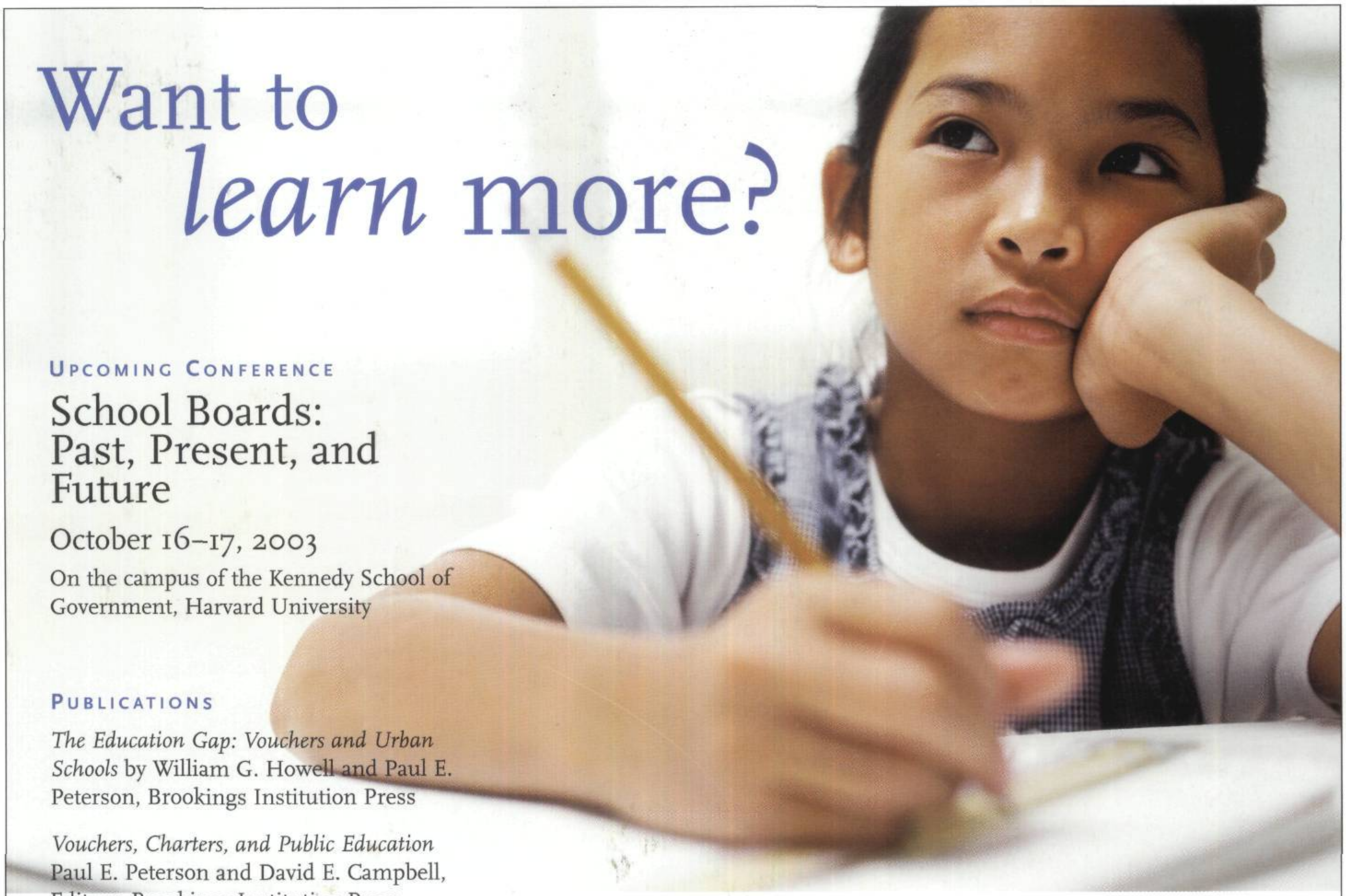
new from HOOVER PRESS

Our Schools & Our Future assesses the changes that have occurred in the twenty years since the National Commission on Excellence in Education urged major reforms in American education. It offers recommendations based on three core principles—accountability, choice, and transparency—that can reinvigorate the system and rekindle America's confidence in its schools.

ISBN 0-8179-3921-0 (cloth) $25.00
ISBN 0-8179-3922-9 (paper) $15.00

to order, call 800.935.2882

HOOVER PRESS
HOOVER INSTITUTION, Stanford University
Stanford, California 94305-6010
Toll-free: 800.935.2882 Fax: 650.723.8626
www.hoover.org hooverpress@hoover.stanford.edu



Want to *learn* more?

UPCOMING CONFERENCE

School Boards: Past, Present, and Future

October 16–17, 2003

On the campus of the Kennedy School of Government, Harvard University

PUBLICATIONS

The Education Gap: Vouchers and Urban Schools by William G. Howell and Paul E. Peterson, Brookings Institution Press

Vouchers, Charters, and Public Education
Paul E. Peterson and David E. Campbell, Editors, Brookings Institution Press

FORTHCOMING

The Future of School Choice
Paul E. Peterson, Editor
Hoover Press

For more see:
www.ksg.harvard.edu/pepg/

HARVARD UNIVERSITY

THE PROGRAM ON EDUCATION POLICY AND GOVERNANCE

Paul E. Peterson, Director
The Taubman Center for State and Local Government
John F. Kennedy School of Government
Harvard University
79 J. F. Kennedy Street, Cambridge, MA 02138
(617) 495-7976 pepg_administrator@ksg.harvard.edu

EDUCATION next
A JOURNAL OF OPINION AND RESEARCH

226 Littauer North Yard
1875 Cambridge Street
Cambridge, MA 02138-3001
Tel. (617) 495-2491, Fax (617) 496-1507
Website www.educationnext.org

Nonprofit
U.S. Postage Paid
Permit No. 26
Peterborough, NH
03458

10071 10070 Xxxx
Ron K Unz
No. 371
555 Bryant St
Palo Alto CA 94301-1704

To subscribe or to submit change of address information, call us at 1-800-935-2882; visit us at www.educationnext.org; fax us at (650) 723-8626; email us at hooverpress@hoover.stanford.edu; or write to *Education Next* Subscriptions, Hoover Institution, Stanford University, Stanford, CA 94305-6010